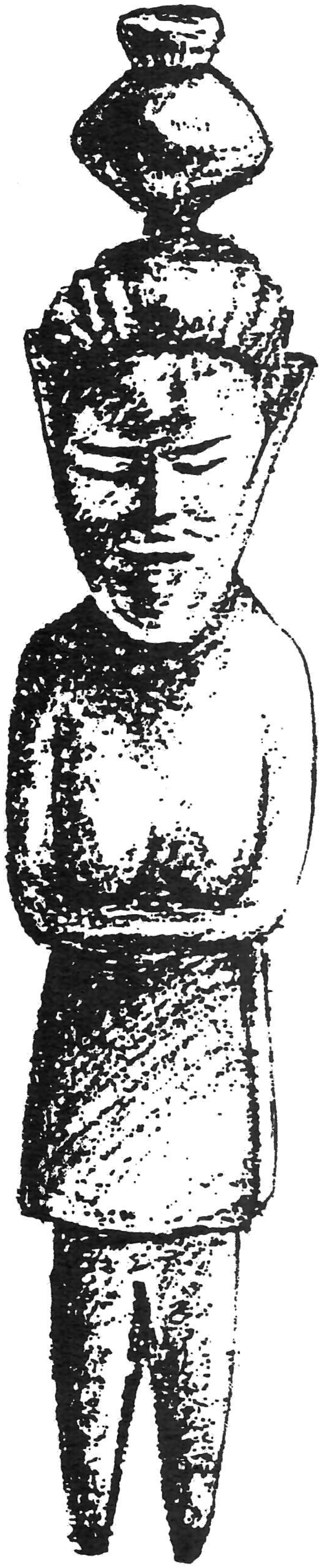
MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY

Issue II, December 1996



CONTENTS

1 Introduction: *John D. Bengtson*

The Nihali Language

2 Special Topic: The Nihali Language
5 Who are the Nihals? What Do They Speak?: *Asha Mundlay*
11 Cognates in the Nihali Lexicon: *Asha Mundlay*
17 Nihali Lexicon: *Asha Mundlay*
41 Nihali Lexicon: Supplement I, Words Collected by Bhattacharya and Konow
45 Nihali Lexicon: Supplement II, Nihali Numerals
46 Nihali Lexicon: Supplement III, Dental and Retroflex Variants

Discussants:

49 Much Ado About Nothing: *Paul K. Benedict*
51 Nihali and Ainu: *John D. Bengtson*
57 Seeking the Relatives of Nihali: *Václav Blažek*
61 Nihali and Nostratic: *Aaron B. Dolgopolsky*
67 Looking to the West and North: Nihali and Kusunda Find Links: *Harold C. Fleming*
75 Nihali and Austroasiatic: *Ilia Peiros*
91 Some Remarks on John Bengtson's Comparison of Ainu and Nihali: *Alexander Vovin*
93 On Nihali: *Norman Zide*

Debate and Discussion: Basque and Dene-Caucasic

101 Comments on the Basque-Dene-Caucasian Comparisons: *Sergei A. Starostin*
111 Response to Starostin: *R. L. Trask*
119 Response to Trask: *Sergei A. Starostin*
125 Second Response to Starostin: *R. L. Trask*
131 Response of Larry Trask to Merritt Ruhlen
137 Letter from Merritt Ruhlen to Larry Trask

Miscellanea

141 Editorial: Comments on Sergei Nikolaev's Letter to Harold C. Fleming (2 September 1995) [*Mother Tongue* (Newsletter) 26:19-22]: *John D. Bengtson*
145 The Genesis of Multilateral Comparison: *Joseph H. Greenberg*
149 References to *Mother Tongue II*

OFFICERS OF ASLIP

(Address appropriate correspondence to each.)

President:
John D. Bengtson
1329 Adams Street NE
Minneapolis, MN 55413
U.S.A.

Vice Presidents:
Daniel McCall
7 Wigglesworth Street
Boston, MA 02114
U.S.A.

Roger W. Wescott
16-A Heritage Crest
Southbury, CT 06488
U.S.A.

Secretary-Treasurer:
Harold C. Fleming
16 Butman Avenue
Gloucester, MA 01930
U.S.A.

Editors of *Mother Tongue*: John D. Bengtson, Daniel McCall, Roger W. Wescott

BOARD OF DIRECTORS

Ofer Bar-Yosef
Harvard University

Anne W. Beaman
Brookline, MA

Allan R. Bomhard
Charleston, SC

Ronald Christensen
Lincoln, MA

Jerold Harmatz
Tufts University

John Hutchison
Boston University

Judith Leader
Lexington, MA

Philip Lieberman
Brown University

COUNCIL OF FELLOWS

Raimo Anttila
UCLA (USA)

Luca Luigi Cavalli-Sforza
Stanford University (USA)

Igor M. Diakonoff
St. Petersburg (Russia)

Aaron Dolgopolsky
University of Haifa (Israel)

Ben Ohiomamhe Elugbe
University of Ibadan (Nigeria)

Joseph H. Greenberg
Stanford University (USA)

Carleton T. Hodge
Indiana University (USA)

Dell Hymes
University of Virginia (USA)

Sydney Lamb
Rice University (USA)

Winfred P. Lehmann
University of Texas (USA)

Karl-Heinrich Menges
University of Vienna (Austria)

Colin Renfrew
Cambridge University (UK)

Vitaly Shevoroshkin
University of Michigan (USA)

Sergei Starostin
Russian State University (Russia)

ISSN: 1087-0326
World Wide Web address: http://www.leonline.com/aslip/index.html

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Introduction to *Mother Tongue II*

Welcome to the second annual issue of *Mother Tongue* (The Journal). The centerpiece of this issue is the *Nihali* (= Nahali, Nehali, Nehari) language of central India, which gets the *MT*Treatment* this time around (see the introduction to *Mother Tongue I*). We are pleased to offer previously unpublished Nihali language data collected in the field by Dr. Asha Mundlay of Pune, India. This will be accompanied by Dr. Mundlay's interpretation of Nihali origins, and finally by the comments of several discussants. Some of the discussants we had hoped for were not able to participate, or did not get their manuscripts ready in time for this issue. Therefore, we expect to continue the discussion of Nihali in *Mother Tongue III*. The references cited by the discussants are combined in a bibliography at the end of the Nihali section.

The other major segment is a sequel to the discussion about the Basque (Euskara) language that was begun in *Mother Tongue I*, and continued in *Mother Tongue* (Newsletter, Spring 1996) 26:19-36. First is Sergei Starostin's contribution to the debate in *Mother Tongue I*, followed by Larry Trask's rejoinder, Starostin's response, and Trask's counter-response. We are pleased to have the comments by Dr. Starostin, one of ASLIP's Council Fellows.

Next we have Dr. Trask's letter to Merritt Ruhlen, responding to the latter's "Reply to Trask" (MT Newsletter 26:22-25), followed by Ruhlen's reply. As in the Nihali section, references are combined at the end.

Finally, as an editorial, I submit my comments on Sergei Nikolayev's letter to Hal Fleming (MT Newsletter 26:19-22).

I am immensely grateful to Allan Bomhard, Hal Fleming, Dan McCall, and Roger Wescott for their assistance in getting this issue of *Mother Tongue* to press.

John D. Bengtson

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Special Topic: The Nihali Language

We usually think of India as a country, but it is really a subcontinent, full of amazingly diverse terrains, cultures, races, and languages (especially when we enlarge the scope to "Greater India," including what are now Pakistan and Bangladesh). As former ASLIP President Hal Fleming reminded me in a recent conversation, Greater India is home to no fewer than *seven* language phyla (including isolates). In generally accepted classifications, these are: Indo-European, Dravidian, Munda (Austroasiatic), Sino-Tibetan, Burushic, Kusunda, and *Nihali*, though as will be seen below, some of us would conjoin some of these phyla to end up with three or four. Even so, it is clear that the Greater Indian subcontinent rivals Africa in linguistic diversity, and as a possible secondary center of human diaspora. It is in this context that we approach the problem of the Nihali language.

Let's put Nihali on the map. (See the map immediately following this introduction.) This map shows, for bearings, the major linguistic boundary of India: the approximate border between Indo-Aryan and Dravidian, bearing in mind that some islands of Dravidian exist to the north and west of that boundary (shown in black on the map), as far as Brahui in Pakistan; and that islands of Munda are found on both sides of the Aryan-Dravidian boundary (see Pinnow's [1959] map).

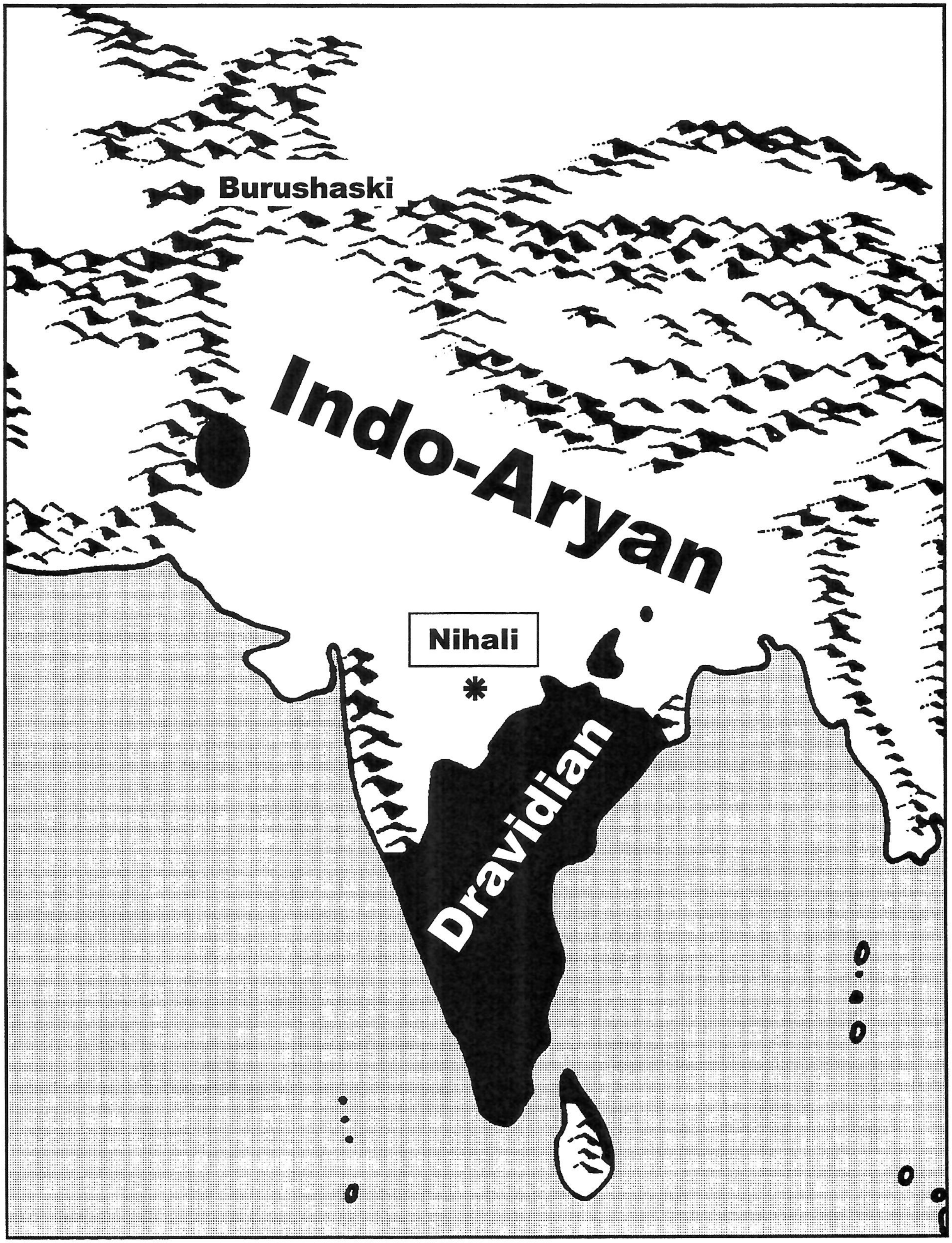
Nihali speakers live in Madhya Pradesh, in the Satpura hills, between Burhanpur and Amravati, and just south of the Tapti River (the location is pinpointed on the map). Norman Zide (Kuiper 1962:243) mentions a village Temi (or Tembi), and Asha Mundlay (personal communication) identifies Sonballi village as the home of Gullu Patel, the sole informant for Sudhibhushan Bhattacharya's (1957) field notes, and Cicari, where she did field work.

The Satpura Hills are about as close to the center of Greater India as one can get, almost equidistant between Gilgit and Trivandrum, Karachi and Calcutta. About 600 miles (1000 km.) to the northeast, in Nepal, we find the few (if any) remaining speakers of Kusunda (Ban Raja), another mysterious isolate; much farther (almost) straight north, about 1000 miles (1650 km.) lies the abode of Burushaski, still generally regarded as an isolate.

"Nihali" is only one of several spellings (see above), but it is the one preferred by Asha Mundlay, and is the form we are using in this journal (except of course when quoting older sources). The Nihals call themselves *Kolṭa* and their language *Kalṭumanḍī*.

According to F. B. J. Kuiper (1962:287) "about 24 per cent. of the Nahali vocabulary has no correspondences whatever in India." Does this "unidentified" lexicon "possibly reflect one of the oldest linguistic strata of India now attainable to research" (ibid., 288)? Is Nihali, as far as can be known, totally isolated? (See Peiros, Zide.) Is Nihali the last remnant of a western outlier of the Austric macro-phylum? (See Bengtson, Blažek.) Are there ties with other macrophyla (Nostratic, Afroasiatic)? (See Dolgopolsky, Fleming.) Or does Nihali, at root, just reflect a divergent branch of Munda? (See Mundlay's article.)

Intrigued by the mystery of Nihali? Read on!


Burushaski
Indo-Aryan
Nihali
Dravidian

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Who are the Nihals? What Do They Speak?

Asha Mundlay
163 / 16 D.P. Road
Pune 411007 India

Telephone: 0212-381099

It is my understanding that the *Kolṭa* (i.e., the Nihals) are an Austro-Asiatic people and speak a language belonging to that family. The branching of that family is still a moot question (see below).

However, it is clear that the genetic connection of the Nihals' speech, which they call *Kalṭumanḍi,* is not very close to that of their immediate neighbors in that family, i.e., the Korku (Kurku) language. Their geographic proximity with the Korkus is historically the result of a deliberate policy on their part to use the peace-loving and land-cultivating Korkus as a camouflage. This was necessary because, in the pre-British era, the Nihals were known as highway robbers, looters of trading caravans, and killers. The Mughal records and the Peshwa records mention this. They had never really been politically subjugated until the British came.

But after 1857, when the British really settled down to rule and administer in a systematic way, the Nihals were no longer in a position to continue their old occupation. In 1870, Captain Forsyth broke the back of many robbing bands, including the Thugs, the Pendharis, and various tribals. (Thugs and Pendharis were not tribals. Their recruitment was voluntary, and from the mainstream population.)

But the tribals (Nihals, Bhils, etc.) had been robbing for generations, and had no interest in agriculture. The Nihals, who had taken to living with the Korkus as marginal inhabitants in Korku settlements in the mid nineteenth century, gradually changed over to petty thieving and raiding of rich men's houses in villages. Some of them also lived as servants of the Korkus. The Korkus felt superior to the Nihals, referring to them as ''hewers of wood and drawers of water," and also as *čomṭa*, which means "thieves." Some Nihals became forest laborers, and some worked as hunting assistants to the British officers.

So for their livelihood, Nihals now have to depend upon contact with their non-Nihali bosses. Most Nihals cannot tell the difference between government employees and the illegal timber contractors, and they work for both with equal zeal, which is to say intermittently and not very hard. Nihals are hired not individually, but as a work group with a leader. These leaders are both men and women, older rather than young, but not necessarily the oldest. The leaders tend to be more intelligent than most of the others, able to keep accounts of wages for the whole group, and know comparatively more Hindi and more Korku than most others.

Here we must look into the reality of linguistic erosion due to inferior social status. Nihals can speak Nihali only amongst each other. I have reason to believe that non-Nihals speaking Nihali are non-existent, except perhaps Sudhibhushan Bhattacharya and myself. Korku, Hindi, and Marathi are the neighboring languages, and their social position is one of dominance. As a result, the present form of the Nihali language shows large-scale borrowing of vocabulary from these sources. Nihali syntax is also influenced by Hindi and Marathi.

The Nihals' neighbors (Korkus and Indo-Aryans) have a prejudicial attitude towards Nihals and their language, based on fear of black magic and thieving skills. Earlier speculations that Nihali is not a real language, but just an "argot" or secret code (see Norman Zide's article), can be explained by this attitude. However, Nihals do have a secret code enabling them to speak in our presence without being worried about our comprehension. This secret speech is falling into disuse as denuded forests and a more thoroughgoing administrative machinery make thieving difficult or even impossible. Urbanization and a changed economic order have destroyed their traditional bonds with "fences," and big robberies have thus stopped altogether.

The Nihali population was decimated in the years 1896-99, due to plague, famine, and a severe shortage of drinking water in their area. An early anthropological map (ca. 1900) shows their number as 18,000. But in the 1981 census, there were only 1,137 Nihali-speaking Nihals, the rest being speakers of Korku. (I have not seen the 1991 census records.) I went to the area in 1988-89 and found them roughly around 1,500 in number.

Why do I believe that the Nihals are an Austro-Asiatic tribe? First, their nomenclature. In the word *Kolṭa* (Nihals), *kol* means 'people' and *-ṭa* is the (Dravidian?) plural suffix. The other Munda words for 'people' are *koro* 'a man' (in Korku), *saora* 'people' (Sora), *ho* 'people' (Ho), etc. Obviously, *kol* is phonologically a cognate of all these terms.

Kal in Nihali means 'woman' or 'wife', and *kalṭumanḍi* means the speech of women. This is consistent with the fact that Nihali women did not have any contact with outsiders and did not speak Korku, but kept to their own speech. Women did not go out to hunt or steal. Interestingly, in Sanskrit, *kalatra* is a neuter noun for 'wife'. The neuter gender shows inferior status: *kal* means 'noise' in Nihali as well as Sanskrit (*kala*, of obscure etymology). The latter could be a very early loanword in Sanskrit, *kalatra* meaning a local (native), noise-making wife, as apart from the regular Indo-Aryan wife, called *patnī*.

Secondly, there is a series of tribe-names spread over a larger area. *Kol* or *kor*, meaning 'people' is used in many place-names in India, all the way from Orissa (e.g., *Korapuṭ*) in the East, to the west coast of Maharashtra (e.g., *Kolapur*, *Kolwaṇ*). The second element is often in a regional dialect (not necessarily Austro-Asiatic) indicating 'place' or 'settlement', or even 'forest'. The northern extension of *kol* place-names has not been investigated by me, but in the South there are *Kolar* (in Karnataka), *Kolam* (in Kerala), and last but not least *Koromandala* (Coromandel) on the southeastern coast. A detailed scrutiny of maps is necessary for an exhaustive list. All these places are villages with a very long history of settlement, but except for *Kolhapur* (in Maharashtra), they are not prominent.

As far north as in Jainsor and Babar, there is an inferior and endogamous caste[1] group, called *Kolṭa*, who look Australoid, very much like our Nihals, who can be described as dark to very dark, but not oily complexioned; with frizzy or curly hair; very dark, even jet-black eyes (as against the usual brown color found in India); the length of the upper torso and the lower torso (from waist to toes) more or less equal; flat-footed; short (around five feet), and stocky in build. This type of appearance is typical of all Munda tribes, and by inference, Nihals must also be related genetically to other Mundas. It is thus a credible theory that they speak a Munda language.

Investigation of local customs, festivals, and gods or goddesses reveals interesting Austro-Asiatic characteristics, e.g.: wooden logs as gods; lack of the concept of afterlife; importance of women after menopause in black magic rituals; nuclear families supported by a complicated system of kin groups sharing meager cash resources; importance of a dancing and singing free-for-all festival in Phalgun (March: the spring equinox), which is the season for matchmaking and sexual freedom, etc.[2]

The presence of Nihals is recognized in our Indian mythology. King Nala in the Mahabharata, who married Damayanti, is believed by the Nihals to be their king. Ravana of the Ramayana is thought by them to be not only their king, but the lord of the place where they go after death. (The euphemism for "died" in Nihali is "went to Ravana.")

Thus, cultural similarities, nomenclature, and names of settlements are the main arguments, outside of linguistic data, that lend credence to the hypothesis of the Munda affiliation of Nihali.

However, the real clinching of the theory must be supported by linguistic data. When borrowings from other languages are taken away, Nihali has its own residual vocabulary, consisting of (a) Munda elements of varying validity (i.e., some borrowed, from Korku, etc.; as well as some native Nihali words), and (b) ''unknown'' words (that is, elsewhere in India). Some examples of (a) are as follows:

39. *alogo* 'strip of bark used as a rope': cf. Parengi *lua:-*, Sora *lua:d-ən* 'fibrous bark used as a string';

185. *bendi* 'jungle': cf. South Munda *bir-*;

205-207. *béthe-* (negative particles): cf. North Munda *ban*, *baŋ*;

229. *biṭini* 'to spread': cf. Gutob-Remo *bed-saq*;

230. *biṭil* 'sand': cf. Gutob-Remo *bitiq* 'salt';

244. *bor* 'hard': cf. Korku *boboṛ*;

343. *cán* 'fish': cf. Korku *kaʔ*;

344. *cakhāw* 'to sweep': cf. South Munda **juq-*, **joq-*;

352. *caré* 'to cut': cf. Korku *cereq* 'cut wood';

354. *cārko* 'black-faced monkey': cf. South Munda *-saq*, *saraq* 'monkey';

399. *cilar-* 'lice': cf. Munda **si-*, **sen-*;

550. *dhonḍum* 'mouth (of an animal)': cf. Parengi *to:t*, Sora *tʔo:d-ən*, *tʔu:d-ən* 'mouth, jaw';

788. *ǰiki* 'eye': cf. Juang *je-* (in *je-tej* 'eye sand');

858. *kalṭo* 'Nihali'; 934. *kol* 'woman, wife'; 935. *kolṭa* 'Nihals' cf. Munda *koro*, *horo*, etc. 'man, people';

944. *kór* 'to take away': cf. Korku *kul/kol* 'send';

1034. *laiñ* 'tongue': cf. Korku *laŋ*, Parengi *-la:ŋ*, Remo *-le:a·ŋ*, etc.;

1118. *minḍi-* (in *minḍi-jiki*) 'wrist, knee' (depending on preceding genitive of 'hand' or 'leg'): cf. Parengi *mandi:*, Sora *mandidi-n*, Gutob *ma:ndi:* 'knee';

All these correspondences with the Munda group are certainly not accidental.

The second part of the linguistic argument is grammatical tendencies, which are similar to Munda language structures. First is the presence of phonemic tone in Nihali. This is not found in the surrounding Indo-Aryan languages, nor does Korku show it in a comparable form. But some other Munda languages do show tone accent alternating with a double consonant.

Another comparable trait is using masculine and feminine personalizers in addressing one's own relatives:

e.g., *baw-re* 'my father'
māy-ja 'my mother'

This *-re* ~ *-ri* is also used for personalizing body parts, intimate possessions, etc.:

e.g., *bhāw-ri* 'my back'
jiki-kāp-ri 'my eyelid'

Then there is reduplication for emphasis: e.g., *helle-helle* 'this much'; *phunṭol-phunṭol* 'baby rattle'; *khoḍḍo-khoḍḍo* 'place behind the knees'.

The known verb system is too impoverished to make any reasonable guesses. But on the basis of stray attested forms, an *Old Nihali (before the famine?) can be imagined.

(1) e.g., *nava oro koʔóyʔ* 'which millet should I bring?' *ko* 'to bring'. Here we find a first person imperative. Also, tone pattern is changed.

(2) verb + *da*: this is used for a verbal adjective or a perfective participle. In this connection cf. in the LSI text *harpi-da* (last but one word in the text).

(3) *ki-ken* for intensification of ordinary *-ken* for future. Cf. Korku intensifiers.

(4) verb + *la*, verb + *kon(a)*: these two are present and future participles, respectively.

In sum, the residual lexical items and the residual morphology are of a distinctive South Munda character. We now know that this South Munda base has been heavily overlaid by early borrowings from Dravidian, later also from Indo-Aryan neighbors; and most recently there has been wholesale borrowing from Korku.

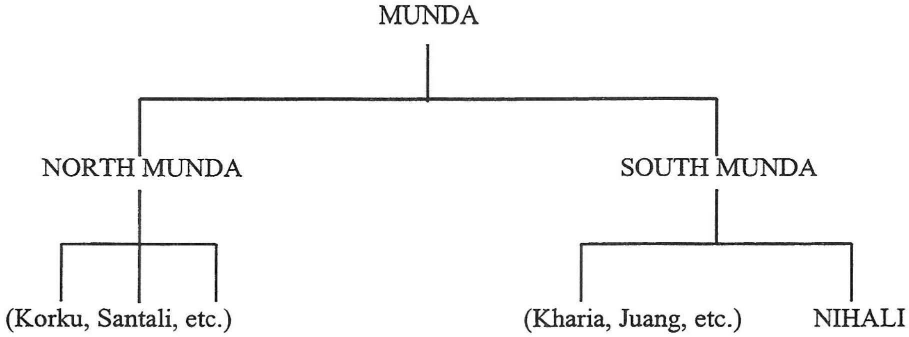
Conclusions

Kolṭa (Nihals as a tribe) and *Koḷi* (a caste in Maharashtra, speaking Marathi) are related. The third possible kin are the *Kolaṭi*, a wandering caste in Maharashtra, who also now speak Marathi, but have some special words used within closed circle communication, and these words and phrases are neither Indo-Aryan nor Dravidian, but are not a code or cipher. They are residual items from a language now long forgotten by them.

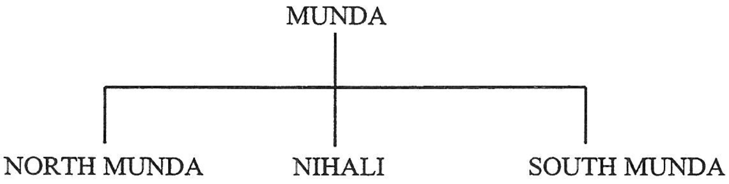
I posit a large group of Austro-Asiatic (Munda) communities over a large part of India. From the west coast to Orissa, all of them are now marginalized. Some have lost their language altogether (e.g., the Bhils); some speak it, but in a much eroded version (e.g., Nihali); and some are still quite alive (e.g., Sora). The linguistic and cultural traits discussed above make me think of them all as one people, and one language family.

My arguments for regarding Nihali as a Munda language follow the axiom that a language is guilty of genetic contact unless proved incontestably innocent of it. It is unlikely that a homogeneous, well-defined group of people, who have Munda-like religious beliefs, food habits, social rules, and a continuous history (since the 1890's) of living in contact with another Munda group (Korku), would have non-Munda linguistic origins. It is to my mind significant that they turned to the Korkus in time of danger, and when faced with extinction. Before the great famines, they lived isolated, but since 1899 they have been camouflaged by the Korkus.

It is possible that the subgrouping is as follows:[3]



Or, it could also be:



But there cannot be a separate language family with only Nihali remaining in India, and hopefully others in Polynesia and elsewhere. That would be far-fetched, and not consistent with lifestyle and cultural patterns. There is no reason to presuppose an altogether different stock.

Notes

1. In my opinion, the Indian people need not be divided into castes and tribes in a rigid fashion. It is better to first map them out in cultural areas which are geographically continuous (as suggested by Massica and Ramanujam of the South Asia Studies Department, University of Chicago). Then, in each area, we divide the people into endogamous communities, some of which are sedentary, and some nomadic, wandering along fixed routes in fixed seasons. Others are sedentary in lifestyle, but migrate once a year to a different locale for economic reasons.

   All of these are referred to as a particular *jāti* (caste) by the common people, defined as an endogamous clan having certain rites in common, certain time-honored customs and beliefs, and a traditional occupation. Subdivision of a *jāti* is possible, but merger is never possible.

   All the *jātis* in India today are clamoring for a separate identity. But they can largely be traced to an original habitat, original linguistic affinity, and acknowledged place on a vertical social status scale. The only real outsiders to this scheme are those who cannot claim such affinity (e.g., foreigners married to Indians, the entire Parsi and Iranian community, and the Indian Jews). On the other hand, Indian Muslims and Indian Christians all have caste/tribe identity, and both demand reservation under the scheduled caste/tribe categories.

   In spite of close intermingling and interdependent living, the Nihals have not merged with the Korkus, and do not intermarry. On the contrary, their religious beliefs actually forbid them to do so. If Nihals were genuine outsiders, linguistically or culturally, they would have shown it. They do not. All they show is insufficient and eroded affinity. So I do not believe in a separate language family for them.

2. Gregory Bateson, in his book *The Naven* (published in 1934), argues that there is a continuous culture area of ritual and music in Polynesia centered around the spring equinox, freedom of mating, choosing marriage partners, singing, dancing, drinking, drum playing, carnivals around a wooden pole (probably symbolizing the male organ). From my knowledge of the music of the *Kol* (Munda) people and their rituals at the *Holi Paurnima* (full moon in Phalgun, i.e., near the spring equinox), I would like to posit a close connection between the Munda people and the Polynesians. At least in the remote past Munda must have come to India from the southeast Asian direction. Similar traditional musical patterns on the drum point to cultural homogeneity in the remote past. [Though the stated "lack of the concept of afterlife" seems to be contradicted by the later reference to Ravana, Dr. Mundlay explains that "went to Ravana" ends there, without any elaboration of the afterlife. Ed.]

3. The "genetic" lineage of a language is only a metaphor. After centuries of constant co-existence with other families, sometimes a language acquires such an overlay of heterogenic characteristics that the genetic picture can become obscure.

   Marathi, my mother tongue, for instance, belongs to the Indo-Aryan family. Yet there is a subtle intermesh in its phrase construction. Influenced by its neighbor Kannada, a Dravidian language, modern Marathi structures cannot be satisfactorily explained without reference to Kannada. The makeup of Nihali, with a very heavy overlay of Korku, Hindi, and some Marathi, is not surprising, given the Nihals' present whereabouts and economic status.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Cognates in the Nihali Lexicon

Asha Mundlay

The composition of the Nihali lexicon is as follows:

1. *Tatsama* borrowings from Korku, e.g., *likhatiŋ* 'to write', where *likh* is Hindi borrowed by Korku plus a Korku verbalizer. [In Indic linguistics, *tatsama* (Sanskrit 'that same') refers to words adopted bodily from another language, usually Sanskrit. The antonym is *tadbhava* ('that be[ing]'), as in *ghām* 'sunshine', which is a regular development of Old Indic *gharma-*, rather than a loanword from Sanskrit. Ed.]
2. *Tatsama* borrowings from neighboring dialects of Marathi and Hindi, i.e., Varhadi and Nimari respectively. It is consistent with the social status of Nihals that the borrowed forms correspond to the lower caste/class dialects of the area; e.g., *ḍhepi* 'wall', *nindo* 'sleep' (Hd.), *ĩway* 'in-law' (Md.), etc.

   Through the medium of the two donor languages, English, Arabic, and Persian words are also taken, e.g., *redigo* (English), *kimto* (Persian), *motar* (English) for 'radio, price, car'. The sound-changes in foreign adaptations are fairly regular. The major part of the lexicon today is composed of 1 and 2 above.
3. Indo-Aryan borrowings which are not *tatsama*. Phonological correspondences here are also fairly regular. Some items which are used in Nihali are not longer found in neighboring Md. and Hd., e.g., *ghām* 'sunshine' (Skt. *gharma-*), *jhatā* 'sister-in-law', *gorši* 'butter-milk', etc. This is important historically.
4. Dravidian borrowings: Here the donor language could be Gondi even though their association was historical and no longer there. But this does not explain all such words. It merits further comparative work, e.g., *ballā* 'stone', *pírjo* 'daughter', etc.
5. Munda (or Austro-Asiatic) elements of varying kinds. These range from recent loan-words from Korku, older loans from other Munda languages, as well as some Munda-like words that may be very ancient, i.e., native to Nihali.
6. Words unknown elsewhere in India, e.g., *cón* or *coʔon* 'nose'.

Now follow some hitherto unpublished lexical resemblances belonging to the fifth class. This is later supplemented by items discussed by Kuiper. Serial numbers refer to the Nihali word-list of 1660 entries.

20. *ajá* 'to thank; please' [ajáa], cf. *ayá* 'to be happy'.

172. *báy* 'today' [báɛ-ɛ̆], [báʔay], cf. SM. Re. *oi*? Ga. *en*.

185. *bendi* 'jungle', cf. SM. *bir*.

209. *bí* 'to rise' [biʔi], cf. K. *bid-*.

229. *bitini* 'to spread', NM. *bil-*, *bel-* 'to spread', GR. *bed-saq*.

230. *bitil* 'sand', cf. GR. *bitiq* 'salt'.

234. *bogon* 'cry of a monkey', cf. SM. *bor-*.

235. *bokki* 'to construct; to tie; to bind', cf. SM. *bod-*. *-ki-* in Nihali is added to verb roots in very few cases. Its function is obscure; but in Korku, *-ki-* is regularly used as an intensifier for verb roots.

241. *bom(m)oki* 'siblings', see below for the word for 'navel'; alternate form *bũmũki*.

244. *bor* 'hard', cf. K. *boboṛ*.

245. *bora* 'scrotum', cf. K. *boḍa* 'penis'.

260. *bulu* 'bubble', cf. SM. *bubu-lañ*.

266. *bumluj* 'navel', cf. K. *bũbluj*. It is interesting that M. *bembi* is 'navel'.

N.B.1. *bakane* 'abandoned', cf. K. *bae* 'leave, abandon', SM. *bañ* (Nihali *bakan* 'to divorce').

299. *b(h)erjoli* 'a small red jungle flower used in cooking', cf. 185 above; *ber-joli* (?).

311. *bhũa* 'trail of a wild animal', cf. SM. *pudaq-*.

?315. *bhúm* 'to move about' [bhúum], cf. SM. *poi-*.

325. *caggi* 'to aim; to shoot'. It could be split as *cag-ki*, cf. Munda **tiñ*.

329. *caka* 'to catch', cf. SM. *cab-/jab-* 'to catch', alternate form, *ceki*.

334. *cakhaw* 'to sweep'; alternate form (1514 *chakaw*). Bhattacharya's form *cokob* for 'leaves'. Cf. SM. **juq-*, **joq-*. Also note Nihali (815) *junu* 'broom; sweep', which may be closer.

338. *calom* ~ *ŋ* 'to beat a drum; thunder', cf. GR. *luŋ-* — but here the semantic considerations leave it moot. *luŋ* 'thank'.

340. *calphẽj* 'covering of a cooking vessel', cf. K. *calpan* 'palm of the hand'.

342. *cana* 'dance', cf. K. *su-sun* 'dance', SM. *su-sun* 'to accompany.

343. *cán* 'fish' [caʔan], cf. K. *kaʔ*.

350. *capir* 'an edible root', cf. SM. *sa-* classifier.

352. *care* 'to cut', cf. K. *cereq* 'cut wood'.

353, *carkhad* 'waist', cf. SM. *ki-sab*.

354. *carko* 'black-faced monkey', cf. SM. *-saq*, *saraq* 'monkey'.

355. *carmaru* 'centipede', cf. SM. *mar-* 'centipede, scorpion, stinging insect'.

374. *cekoṭo* 'to chop meat', cf. GU. *seb-* 'slaughter'.

386. *cergo* 'run', cf. K. *saṛub-*.

391. *cicidoke* 'a species of birds', cf. K. *tid-* 'bird'.

395. *cikin* 'worm, small winged insect', cf. Munda *si/se* 'louse'. Also *cilar* 'lice'. The latter may be a closer parallel.

455. *dak(a)ri* 'old, aged', cf. Re. Gu. *dukri*, *dokri* 'old'. Rajasthani and Gujarati dialects also show *dokra* 'old man'.

465. *dangi* ~ *dango* 'branch of a tree', cf. Sora *daŋgū-n*, Gutob *daŋgu*.

512. *dubumbi* 'abdomen', see above *bumlui* 'navel' and its cognates.

833. *kajito* 'thief' (historically, Nihals killed when stealing), cf. Parengi *oggi*, Sora *ab-kajeḍ*, *akkajeḍ*, *kajjeḍ*.

517. *dukrimay* 'father's older sister; grandmother', cf. Sora *dukkirīn*, Remo *dukri* 'wife'.

1101. *mengé* 'tooth/teeth', cf. Parengi *dyi*, Sora *ɟi*, Remo *giŋe*, Gutob *giŋe*.

1026. *laiñ* 'tongue', cf. K. *lan*, NM. *alaŋ*, Parengi *-lāŋ*, Sora *la·ŋ-ən*, Remo *-le:a·ŋ*, Gutob *-lāŋ*. Here it means 'tongue, language'.

1032. *laŋka* 'far-away place', cf. Sora *laŋka*, GR. *duŋku* 'high'.

L1385. *raban* ~ *roban* 'cold', cf. K. *raban*, Sora *vaŋān*, GR. *ruoɲ*.

550. *dhonḍum* 'mouth of an animal', cf. Parengi *tōt*, Sora *tʔōd-ən*, *tʔūd-ən* 'mouth, jaw'. Sanskrit *tuṇḍam* is a relation?

470. *darom* 'to escort', cf. Sora *u'ruŋ*, *o'roŋ*, Remo *riŋ*, Gutob *indriŋ*, *giriŋ*.

39. *alogo* 'strip of bark used as a rope', cf. Parengi *luā-*, Sora *luād-ən* 'fibrous bark used as a string'.

133. *minḍi-jiki* 'wrist, knee' (depending on preceding genitive of hand, leg), cf. Parengi *mandī*, Sora *mandidi-n*, Gutob *māndī* 'knee'.

137/138. *bakí* ~ *baka* 'to trap', cf. K. *baq*, *-k-* may be separated from the root.

141. *bamaka* ~ *bamaki* 'otherwise', cf. K. *ban-ba-ken*, lit. 'if not, then'.

157. *báro* 'sing, song', cf. SM. *bir-ʔ*.

?160. *baruʔu* 'to bite', cf. SM. *bʔe* 'to eat a lot'.

164. *base kol* 'younger wife', cf. SM. *sel-* 'woman'.

?186. *benge aji* 'husband's elder sister', cf. SM. *b-ən*.

?194. *beri* 'to beat a drum', cf. Sk. *bheri* 'one drum in war'. But SM. *bad-*, *bug-*.

200. *beso* 'to whistle', *bešo*, *becho*, cf. GR. *sɨg*.

202. *bétki* 'to sow', cf. K. *bid-*. Alternate form *bekki* (in Nihali); *-ki-* can be separated.

297. *bethela* ~ *biji* ~ *biʔ* ~ *betheʔy* negative particles, cf. NM. *ban-baŋ*.

217. *bigi* 'to blow (on a fire)', also 'to play flute by blowing'. It could be *bi -ki*, cf. SM. *ped-* 'to blow a musical instrument'.

240. *bombo* 'cry of anguish at childbirth', cf. M. *bomb* 'cry loudly' but also Re. *bum-bo* 'sexual intercourse'.

249. *capiniku* Nihali name for Korkus. This may be a case of infix: *ca - pi - ni -ku*
infix pl.
cani ~ *sani* 'small'.

?416. *có* 'urinate', cf. Nihali *cópo* 'salt'.

?422. *comṭa* derogatory term of reference by Korkus and Nihalis for each other; *com-ta* (plural).

434. *cucormugũĩ(j)* 'house lizard', cf. K. *kekhemered*.

558. *dhuk-dhuk* 'to beat (of a heart)', cf. SM. *duq-*, but also M. *dhuk-dhuk*.

598. *gánda* 'to split wood', cf. *ganduʔi*.

L758. *jádi* 'root' [ǰáṛi], cf. Sora *'jʔēd-ən*. But it may be Indo-Aryan.

739. *iŋgin* 'we' (pl.). Since pronouns have dual forms in Nihali, *iŋgin* may be connected with Munda numeral 'three', cf. GR.

788. *jiki* 'eye', cf. Juang *je-tej* 'eye sand'. To isolate *je-tej*, cf. *mo-tej* 'nasal mucus'.
819. *jhallya* 'peacock', in Ga. 'male peacock' *jhaḷia*.
873. *kaplij* 'butterfly', cf. K. *kaphiliq*.
?884. *karjo* 'cashew'.
887. *karwẽñ* this is an alternate form of *kakhẽñ* 'comb', see below.
889. *katam* 'silence', cf. K. *katan*.
897. *kat(h)la* 'armpit', cf. Mundari *hatalaʔ*.
900. *katú* 'embrace', cf. SM. *kũdu*.
934. *kol* 'woman, wife', cf. *koro* 'man'; also its cognates *horo*, etc. The Nihali name for themselves is *kolṭa/kalṭu* 'men'.
?944. *kór* 'to bring away', cf. K. *kul/kol* 'send', SM. *kir/kur* 'to roll'.
1076. *mando* 'frog' and 1077 *mándu* 'rain' are puzzling, but less so when cf. Sk. *manduka* 'frog' (who traditionally cries at advance of rain). This has been thought of as coming into Sk. from Proto-Munda.
1105. *meur* 'anthill', cf. SM. *reor* and K. *bunum*.
1203. *oda* 'temple', cf. NM. same for 'house'.
1524. *chemic* 'chicken dung', cf. *siñ*, etc. for first part (SM); *ic* can be related to *ij*, Korku for 'excreta'.
1633. *ujori* 'light (as opposed to darkness)', cf. *ujun*, etc. in RG. for 'sun'.

Kuiper has previously pointed out the following items which I have listed here separately:

1. *biḍiko* 'one'. Kuiper gives *biḍi*, but alternants *biḍiko*, *biḍi*, *baḍa* are found in Nihali. Its cognates either begin with 'b' without the dental suffix or with 'm' plus or minus suffix.
2. *té* 'eat', cf. *tin kherwani*. Sa. **aṭiń* 'to graze'.
3. *chama-ki*, cf. Korku **samah*. Zide comments that this is contaminated by Hindi *samne*.
4. *hondar* 'rat'; this resembles Proto-Munda, which is reflected in Sanskrit as *undara*.
5. *haru* 'to bite'; Kherwani **haru*.
6. *batam* 'thirst', cf. K. *tatan*.
7. *ará* 'to see'; Sa. *arak*, *orok* 'staring'.
8. *giṭa* 'younger brother', cf. Sa. *giḍar*, *gaḍar* 'little children'.
9. *ho-ṭ*, *ho-ṭe* 'not', cf. Sa. *o̱ho̱*.
10. *jakoṭo* 'male', cf. Kw. *jhāku*.
11. *mokhne* 'elbow', cf. M. *kukũri* 'knee'.
12. *popo* 'belly', cf. M. *pu'pu'*, Ho. *pupū*.
13. *kakhẽñ* 'to comb', cf. K. *akhej*.
14. *arḍu* 'tree' (*aḍḍo* in my list), cf. So. *ĕrā*?
15. *baddi* 'bull', cf. Gu. *bāḍi* 'buffalo'.
16. *be* 'to give', cf. Gu. *bē*.
17. *de* 'to give', cf. Ju. *din*. This is suspect. H. *dena*?
18. *ēr*, *iēr* 'to go', cf. So. *er*, *yir-*.
19. *jere* 'to remain', cf. Ju. *iḍ*, *iṛ*?
20. *piy* 'to come', cf. Gu. *pi*. This to my mind is *bí* in Nihali with *bíʔiya* 'past'.
21. *aba-re* 'father'. I do not accept this, as *aba* has obvious Dravidian cognates *aba*, *apa* for 'father'.

The following correspondences emerge from perusal of the above:

Nihali	South Munda	North Munda
b	b ~ m	b ~ p
bh	p	?
c	j	t
c	c	?
c	s	k
k	k	k
d	d	
g	g ~ j	

Nihali	South Munda	North Munda
l	l	l
l	r	
r	r	
dh	t	
m	m	
dh	dh	d

The picture of vowels is unfortunately obscure, especially because of tone and its neutralization in Nihali, but:

Nihali	South Munda	North Munda
a ~ ə	a	a
i ~ e	i ~ e	i ~ e
u ~ o	u ~ o	u ~ o

can be seen with slightly less regularity than that in consonants in words accented in Nihali.

It will be clear from the above list that all these correspondences with the Southern Munda group are certainly not accidental. It should also be borne in mind that, even if all tatsama Korku words have been excluded, these may be replacements of Nihali items which were connected with other Munda items. By computation, a certain weight can be given to this assumption.

The second part of the argument is grammatical tendencies which are similar to Munda language structures.

First, the presence of phonemic tone. It is not found in other Indo-Aryan languages surrounding it. Korku does not show it either in a comparable form. But Munda languages do show the accent alternating with double consonant.

In the dual suffixes, cf. *ko* ~ *kel* of the 1st, 2nd, and 3rd person. Dual is now no longer present in nouns excepting in a couple of cases. One is *kalaṭ-ṭel* 'two wives'; *k* > *ṭ* by assimilation. But for other nouns, the numeral *yir* 'two' is prefixed to indicate a group of two.

Here 'k' for non-singular is comparable to what is found in other Munda languages. In the demonstrative also, *-kel* is for non-singular.

Another comparable trait is using masculine for feminine personalizers in addressing one's own relatives:

e.g. *bav-re* 'my father'
may-ja 'my mother'.

This *-re* ~ *-ri* is also used for personalizing body-parts, intimate possessions, etc.

e.g. *bhāw-ri* 'my back'.

Then there is reduplication for emphasis:

e.g. *helle-helle* 'this much'
phunṭol-phunṭol 'baby rattle'
khoḍḍo-khoḍḍo 'place behind the knees'.

The known verb system is too impoverished to make any reasonable guesses.

But on the basis of stray attested forms, an Old Nihali (before the famine?) can be imagined:

1. e.g., *nava oro koʔóy*? 'which millet may I bring?', *kó* 'to bring'. But here we find a 1st person imperative. Also, tone pattern is changed.
2. verb + *da*: this is used for verbal adjective or perfect participle. In this connection, cf. LSI text *harpi-da* (last but one word in the text).
3. *ki-ken* for intensifier of ordinary *-ken* for future; cf. Korku intensifiers.
4. verb + *la*, verb + *kon(a)*; these two are present and future participles respectively.

Nominals

Gender is not now present, but there are words for a pair:

e.g. *getháy* 'a dog and a bitch';

words are also known for 'a cock and a hen' and 'a cow and a bull', but not universally. At present, *jakoto* 'male' and *kol* 'female' are added to animal names to form required compounds:

e.g. *jakoto bérko* 'male cat'.

Not much can be said about grammatical evidence except in a negative fashion, i.e., no known bit contradicts a Munda connection.

The third, or religious-cultural, part has already been dealt with elsewhere. Here again, the argument is based on general patterning in the area. It is true that there is another group, i.e., the Bhils, in the area where linguistic affinities are undecided, but religious practices make Nihals distant from them.

Conclusion

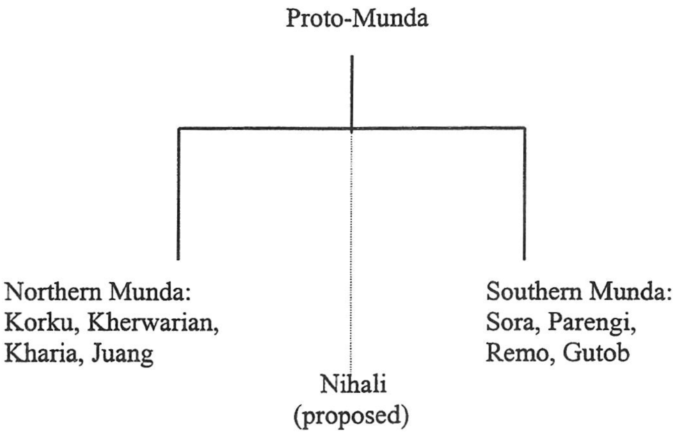
There is thus a prima facie case for classifying Nihali as a Munda language: on lexical, grammatical, and ethnographic evidence. Further research is indispensable to arrive at a more definite appreciation of the position.

However, there still remains the very important question of explaining the geographical discontinuity of Southern Munda languages and Nihali. For the intervening Munda languages belong to the Northern branch. The pan-Indian features of linguistic groupings in India, regardless of language family, do show a bias in a different language than this case. Besides, all historical evidence of Nihali plundering is even westwards of the present position.

Nihali also shows some influence of NM, especially Korku.

And then again, the resemblances do not cover all aspects of Nihali grammar. There are lexical and structural characteristics which are unlike every other Munda language that we know so far. The picture is obscured by structural erosion and large scale lexical borrowing. Nor are the resemblances close enough.

It will thus be safe to leave Nihali out of both NM and SM branches. I present below a simplified sketch of my guess about Nihali in the scheme of Munda languages.



I refrain from elaborating on NM and SM branches. Their sub-divisions have been proposed by Norman H. Zide. I see no obvious reason to dispute his scheme. But I do not reproduce it here.

My arguments for regarding Nihali as a Munda language follows the axiom that the language is guilty of genetic contact unless proved incontestably innocent of it.

It is unlikely that a homogeneous, well-defined group of people, who have Munda-like religious beliefs, food-habits, social rules, and a continuous history (since 1890's) of living in contact with another Munda group would have non-Munda linguistic origins. It is to my mind significant that they turned to Korkus in time of danger and when faced with extinction. Before the great famines, they lived in isolation. After the 1890's or, to be precise, after 1899, they have been camouflaged by Korkus.

Considering how slow language change generally is, the time is not sufficient for Nihali to have undergone a linguistic overhaul from a non-Munda language to its present form, which looks like an eroded Munda language, especially since all language contact of Nihals with others is oral and direct, and the social mobility of Nihals is zero.

The above has been placed as a preliminary statement for consideration by all those linguists who are working on Austro-Asiatic languages. I am well aware of my limitations in linguistic theory and acquaintance with other Munda languages in India.

It is presented with a hope that Nihali comes out of data-oblivion. I have never been happy that endless discussion should go on based only on LSI and Bhattacharya's *Notes* as the sole corpus for investigation.

Abbreviations used in the "Nihali Lexicon"

Adj	adjective	N	noun
GR	Gutob-Remo (SM)	NC	noun compound
H	Hindi	NM	North Munda
Hd	dialect of Hindi	num	numeral
I	indeclinable	NV	nominal-verbal
K	Korku (NM)	Sk	Sanskrit
L	loan-word	SM	South Munda
M	Marathi	Vintr	verb, intransitive
Md	dialect of Marathi	Vtr	verb, transitive

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali Lexicon

Asha Mundlay

a, ā, á

L1. *abā* (N) ‘father’, cf. M. *ābā* term of address, Md. *bāp* ‘father’. Nihali makes no distinction here between term of reference and of address, K. *-ba*, PM. *-ba*.
  a. *abāna bhaga may* (NC) ‘great grandmother’, lit. ‘father’s grandmother’.
  b. *abāna bhaga ābā* (NC) ‘great grandfather’.
L2. *ačambo* (N) ‘surprise’, cf. M. *acamba.*
L3. *ačar* (N) ‘end of a sari which is draped over the shoulders’, cf. Hd. *acal.*
4. *ačči* (Vtr) ‘dig’
L5. *āḍā* (Adj -Vintr-) ‘half; to halve’, cf. H. *ādhā.*
L6. *aḍar* (NVtr-) ‘order’, cf. English *order* in use in H. and M.
?7. *aḍawḍi* (I) ‘suddenly, at once’. This is possibly a reduplication *aḍa-oḍi*. M. *ekaeki* has the same meaning and similar formation, but in M. *aḍ-* ‘to obstruct’, so no connection.
8. *aḍḍo* (N) ‘tree’.
  a. *aḍḍo umúni* (NC) ‘seed bed’.
  b. *aḍḍone joppó* (NC) ‘sap of the tree’, lit. ‘water by the tree’.
9. *aḍé* (Vintr) ‘to burn’ as in ‘the fire’.
L10. *adito* (NV-) ‘habit’, cf. H. *ādat.*
L11. *adla-bedli -kamay* (V) ‘to barter, exchange’, cf. M. *adlabadal* (noun).
12. *admosu* (Vintr) ‘to stretch after sleep.
?13. *ador* (Vintr) ‘to speak haltingly, with a defect’, cf. M. *aḍ-* ‘obstruct’ (verb).
L14. *aggari* (N-) ‘train’, cf. M. *āggādi.*
L15. *aggaso* (N-) ‘sky’, cf. M. *ākāš.*
L16. *agin* (N) ‘summer’, cf. MH *agni* ‘fire’.
17. *agin-bí* (VC) ‘to perspire’.
18. *agor* (N) ‘trap for catching animals in the forest’.
19. *agrí* (Vtr) ‘to shut, close’.
20. *ajá* (V) ‘to thank’, cf. K. *aya* ‘to be happy’.
L21. *ajāb-kamay* (V) ‘decorate’, cf. HM *ajab* ‘wonderful’.
L22. *ajānā* (N) ‘name of a tropical tree’, cf. HM *anǰan.*
23. *āji* (N) ‘sister-in-law’, i.e., ‘husband’s younger sister’.
L24. *ajiw* (I) ‘immediately, now’, cf. Md *ajji* meaning ‘presently’ (said to be from Urdu, but cf. Sanskrit *adya-* ‘today’).
L25. *ajgar* (N) ‘python’, cf. HM *ajgar.*
26. *ajmí* (NV) ‘vomit’.
27. *ákā* (V) ‘to hang up on a line, etc.’, cf. K. *akhà-.*
28. *akāwā* (N) ‘name of a bush’.
?29. *akhaṇḍā* (N) ‘finger’, cf. M. *āŋgaṭhā* ‘thumb’, Sanskrit *anguṣṭha-* ‘thumb’. This must be an older loan since there is metathesis of the nasal and aspiration as well as change in voicing of the consonants.
  a. *sendā akhaṇḍā* (NC) ‘thumb’.
  b. *akhaṇḍā langí* (NC) ‘web of a finger’.
?30. *akhuṇḍa* (N) some speakers return this for 29 above; cf. Sk. *anguṣṭha-*, OM. *anguṭhā.*
?31. *akiri* (N) ‘camp, short-term resting place’, H. *ākhādā* ‘gymnasium of wrestlers’.
L32. *akkal-kāmāy* (VC) ‘to advise’, cf. M. *akkal* (noun) ‘wisdom’ (from Arabic).
L33. *akkal-kāyni* (VC) ‘to cry loudly in anguish’.
L35. *alā* (I) ‘exclamation of surprise’, cf. K. *ala.*
?36. *alābalā* (I) ‘descriptive of inferior quality’, e.g. *alābalā biji bárobe* ‘do not sing such trash’, cf. M. *alābalā* used in the ritual of warding off the evil eye.
L37. *alāgat* (I) ‘descriptive of caution, delicacy, discretion, etc.’, cf. M. *algat* of similar meaning.
L38. *alāgo* (NV) ‘separate’, cf. MH. *alag* ‘separate’.
39. *alogo* (N) ‘a strip of bark used as a rope; a thin bamboo stick’.
40. *alú* (V) ‘to stretch oneself’, cf. M. *āḷəs deṇe* same meaning, Sk. *ālasya* ‘laziness’.

L41. *āmuso* (N) 'new moon night', cf. Md. *amuša*, Sk. *amāvāsyā*.

L42, *anā* or *ānā* ? (N) 'one anna, 1/16 of a rupee'. This was the coin before the decimal system was introduced in India, cf. M. *āṇa*. Now it is used to indicate percentage of crops, etc.

L43. *ana-biki-kayni* (VC) 'to take a vow, to promise secretly', cf. M. *āṇbhāk* (noun) 'secret pact' — *āṇ* refers to the gesture and *bhāk* to the accompanying words; cf. Sk. *bhaṣ* 'speak'.

44. *ancho* (V) 'to pick'.

45. *anchi* (V) 'to select, to pick'; see 44.

46. *anḍa* (Adj - V) 'bad, false; to become bad'.
   a. *anḍa bai* (NC) 'female adulterer'.
   b. *anḍa manso* (NC) 'male adulterer'.

47. *andáy* (V) 'to suffer from venereal disease'. (In M. *ānḍa* refers to a man's balls). Sk. *anḍa* 'egg'.

48. *āndij* (N) 'a root like the sweet potato'.

L49. *andori* (NV) 'darkness; to become dark', cf. M. *andhār*.

?50. *andhloy* (N) 'menstruation'. It could be a NC *anḍa loy* 'bad blood', cf. H. *loi* 'blood'.

?51. *ane* (Adj - V) 'other; to do such and such', e.g. *jó téka do ane-aneka do kappi* 'I eat and do this and that and slept', cf. Sk. *anya*.

52. *anjer* (V) 'to become senseless; faint'.

L53. *angārā* (N) 'fragrant black powder of magico-religious potency', cf. M. *angārā́*.

L54. *angulij* (V - N) 'to bathe; bath' (it may be body-wash). In the nominal form, /j/ is optionally deleted, cf. K. *ãguluj*, M. *ānghoḷ*, Md. *anguli*, Sk. *anga* 'body'.
   a. *angloy basonki bethey* 'puberty', see 50, lit. 'not taking bath'

L55. *angúb* (NV) 'to yawn', cf. K. *angub*.

L56. *angārko* (N) 'upper garment of a man', cf. M. *angarkhā*.

L57. *anuij* (V) 'to bend down (as of a branch)', K. *anuj*.

L58. *āp* (PN) honorific pronoun for second person; used only to address officials, etc., who are non-Nihals; never used in an all Nihali talking group, cf. H. *āp*, M. *āpaṇ*.
   a. *āpla* 'your', cf. M. *āpla* 'genitive form', usage as above only for non-Nihals.

59. *ápa* (V) 'to cry'.

60. *ápaenkāmā* (V) 'to make to cry', e.g. *jó bacakan ápaenkamaka* 'I cause the child to cry'.

61. *aparā* (N) 'ritual impurity after death; mourning', see 59. This could be *apa-ra*; *ra* is a personalizer, lit. 'crying oneself'.

L62. *aphír* (V) 'fly', K. *aphir*.

63. *aphiri* (N) 'fire-fly'.

64. *apó* (V) 'to start towards'.

65. *ápó* (NV) 'wood; to be lit', e.g. *apóka* 'it burns'.

66. *ápokama* (V) 'to light a fire'.

?67. *apse-bando* (VC) 'to check flow of blood' (as in dressing a wound), *bandh* 'to tie' in M.

68. *ara* see (a) below; does not occur by itself.
   a. *ara jā́to* (NC) 'other caste', e.g. *jó kaltu āp babinye. ara jato āpla* 'I (am) a Nihal. You are a brahmin. Your caste is different'.

69. *ará* (V) 'to examine closely'.

L70. *arado* (NV) 'to systematically cry loudly to raise a wild animal in hunting' — this may be an older loan. Hunting by Nihals is now non-existent. Cf. M. *arḍa-oraḍ* (noun) 'loud cry'.

71. *arāngá* (Adj - V) 'stale; to become stale', e.g. *sokora arangaya* 'bread became state'.

72. *arāni* see (a); does not occur by itself.
   a. *arāni mór* (NC) 'a thorny bush, grown for hedges', M. name for it is *ṭaṇṭaṇi*.

73. *arāṭi* (N) 'thorny bush'.

L74. *arāy* (num) 'two and a half', cf. H. *aḍhay*, Hd. *aray*.

75. *arāy* (V) 'to see', cf. 69 above (use of tone by Nihals — some speakers lapse from it)

76. *arec* (V) 'to brush away into a pile on one side (dry leaves, etc.)'.

77. *arki* (V) 'to weave'. Nihals nowadays buy all clothing in weekly markets. But stories of weaving survive in folk tales.

78. *ārno* (V) 'to leave something behind one, when leaving a place'. This has religious significance for them.

79. *ārom* (V) 'to break, fall off (as of teeth)', e.g. *pār aromi mengé* 'all teeth are gone'.

80. *arom-kāmāy* (VC) 'to demolish'.

?81. *aroyní* (NV) 'to respond to a call coming from a distance', cf. M. *aroḷi* 'a loud cry'.

82. *árthi* (V) 'to cause to weep'.

L83. *āṭa* (N) 'food', cf. HM *āta* 'wheat flour', Kannaḍa *uṭa* 'meal'.

84. *āta* (V) 'to distribute (food, etc.)', see 83.

85. *ató* (V) 'to break (a stick, etc.)'.

86. *áto* (V) 'to twist'.

87. *asudi* (V) 'to hang something on a line'.

?88. *asud* (V) 'to be flat on one's back in wrestling'.

89. *átho* (N) 'husband'.

90. *athú* (V) 'to crack the finger joints, etc.', see 85.

91. *attú* (V) 'to stretch', see 85, 90.

92. *ãto* (N) 'tear (of the eye)', cf. 793.

L93. *āwār* (N) 'house', cf. M. *āvār* 'compound area of a house', Hd. *āvār* 'house'.

a. *āwār-nitti* (VC) 'wife who was formerly someone else's wife' — either a divorcee or a widow who is remarried.

b. *āwar-āthoka-bethey* (NC) 'spinster'.

L94. *awol* (Adj - V) 'good', cf. HM. *avval* 'of high quality'.

95. *awoli* (V) 'to keep away from'.

96. *awolo* (V) 'to cure', see 94.

L97. *āy* (N) 'mother', cf. Md. *āy*.

98. *ayi* (I) 'below', used as a pp. also.

99. *ayin-te* (I) 'gently'.

100. *ayin* (V) 'to smell'.

b

L101. *bābā* (N) 'paddy', cf. K. *baba*.

L102. *bābā* (N) 'mother's father', cf. Md. *bābā* 'grandfather on father's side'. This then is a loan with a semantic change.

103. *bābarā* (N) 'an edible root'.

?104. *bābine* (N) 'priest for marriages; a brahmin', cf. Md. *bāman*.

L105. *bābur* (N) 'a thorny tree', cf. Md. *babhur*, M. *babhuḷ*.

L106. *bacān* (N) 'abstaining, refraining', cf. H. *bačānā* (verb), M. *bacauṇe*.

L107. *bač(a)kā* (N) 'small boy, young one of an animal', cf. H. *bacca*. *-ka* is a diminutive Nihali suffix.

a. *bačakā-re* (N) 'one's own child'. *-re* is a personalizer suffix.

b. *bačka-ri* (N) dialect variant of 107a.

c. *bačkā kol* (NC) 'female dwarf'.

d. *backā pasārā* (N) 'narrow passage' (by implication 'narrow cervix' — used when describing child-birth difficulties).

108. *baḍā* (num) 'one'; see *biḍiko* below, also 'identical, alike'. When meaning 'alike', it is used as a nominal suffix.

109. *badār* (V) 'to scratch one's skin', cf. M. *bhadarṇe* 'to shear, scrape'.

L110. *badmāšo* (N) 'criminal', cf. H. *badmāš*.

L111. *bādo* (NV) 'of a later origin', cf. H. *bād* 'later' — notice change of form class from a pp. In H. to NV in Nihali.

112. *baḍágo* (N) 'guava'.

L113. *badlā(y)* (NV) 'revenge', cf. HM *badlā*.

L114. *badrā* (N) 'cloud, sky', cf. Hd. *badri*, H. *bādal*.

115. *bagár* (N) 'young one of a buffalo'.

L116. *bage-bage* (I) 'quickly', cf. K. *bage bage*, Md. *bigi-bigi*, which means 'quickly' but is not explained in M. Marathi could have borrowed it.

L117. *bagičā* (N) 'garden', cf. H. *bagicā*.

L118. *bahirā* (N) 'deaf person', cf. M. *bahirā*.

119. *bahirā-čigám* (NC) 'deaf mute'.

L120. *bāi* (N) 'woman, elder sister', cf. M. *bāi*.

L121. *baika* (N) 'woman, sister', cf. M. *bayko*, which means 'wife, woman'.

L122. *bajā* (I) denoting time by the clock, e.g. *moṭh baja* 'three o'clock', cf. H. *bajā*, *baje*.

L123. *bajān-doyra* (NC) 'religious singing by the Muslims and Hindus', cf. H. *bhajan*, *dohra*.

123/2. *bajo* (N) 'arrow'. (Inserted later.)

L124. *bākā* (V) 'famous', cf. M. *bākā* 'brave'.

125. *baká* (V) 'to scare away'; see 128.

126. *bakān* (V) 'to choke (on food, etc.)'.

127. *bakán*^{1} (V) 'to stick to the ground (an incense stick to appease gods)'.

128. *bakán*^{2} (V) 'to leave, release, divorce, abandon, set free'. N.B. 1: *bakāne* obscure verbal form 'abandoned', cf. K. *bae* 'leave, abandon', SM *baɲ*.

L129. *bakārā* (N) 'he-goat', cf. M. *bakṛā*.

130. *bakrā-kāv* (N) 'lamb curry'.

131. *bakko* (N) 'palm of the hand'.

132. *bakko-koḍḍu-akhaṇḍa* (NC) 'thumb'.

133. *bakko-ko-mindi-jiki* (NC) 'wrist'.

134. *bakko-tepre* (NC) 'knuckles, dorsum of hand', *re* could be personalizer for body-parts.

L135. *bakru* (N) 'kid'; see 129.
?136. *bakāy* (V) 'to bear fruit; to be plentiful; to be in abundance'.
137. *báki* (NV) 'trap; cheat'.
138. *báka* (NV) dialect variant of 137, cf. K. *bag*.
139. *baki-jeré* (V) 'to get caught (on a fish-hook, etc.)'.
L140. *ballā* (N) 'hill', cf. K. *balla*. This could be of Dravidian origin — *ballā* is 'stone/rock' in D.
a. *ballā-kajǎrki kheto* (NC) 'hill field'.
b. *ballāṭa bāju* (NC) 'hillside'.
141. *bamāka* (I) 'otherwise', cf. K. *ban-ba-ken*, lit. 'if not, then'.
142. *bamāki* (I) dialect variant of 141.
143. *bambay* (N) 'Bombay'. Proper name of the big city known to Nihals.
144. *bambogo* (Adj) 'pointed'.
145. *bamrā* (N) 'wick'.
146. *bambārā* (N) dialect variant of 145.
L147. *bana* (N) 'clothes', cf. Md. *bāṇ*, M. *vāṇ*.
L148. *bāp* (N) 'father', cf. M. *bāp*.
149. *baure* (N) 'one's own father'; *-re* is personalizer suffix for kinship.
L150. *baṛā*^{1} (Adj) 'big', cf. H. *baḍa*.
151. *baṛā*^{2} (N) 'threshing floor'.
L152. *barābor*^{1} (Adj - V) 'approve, correct, to be equal to', cf. M. *barobar*.
L153. *barābor*^{2} (pp) 'together with', cf. M. *barobar*.
154. *baṛa joppó-ji* (N) 'wind-pipe', which, according to them, is used for air and water intake; see 155.
155. *baṛa té-ji* (N) 'esophagus'.
156. *bardo* (N) 'sickle'.
157. *báro* (NV) 'sing; a song'.
L158. *barsādo* (N) 'rain', cf. H. *barsāt*.
L a. *barsādo dino* (NC) 'wet season'.
159. *baru* (N) 'mulberry'.
160. *barú* (V) 'to bite'.
161. *basin* (Adj) 'dangerous'.
?162. *base* (Adj) 'small'.
?163. *basel aḍḍo* (NC) 'shrub'.
?164. *base kol* (NC) 'younger wife'.
?165. *base la* (Adj) 'little, small'.
?166. *basel* (Adj) 'many' in one dialect. This is very interesting. It could be ascribed to informant's inadequate grasp of his own language. This phenomenon is quite common among younger Nihals.
167. *basonki*: see 54a.
168. *batāme* (NV) 'thirst; to be thirsty', cf. K. *tam-*, NM *tetaŋ*.
169. *baṭiṭa* (Npl) 'big flies (houseflies)'. The idea of using it in a singular form was laughed at as unrealistic.
L170. *bātuko* (N) 'mango', cf. Md. *baṭhi*, *bitki* for 'mango-stone'.
L171. *bawan* (N) 'wife's sister', cf. M. *bahiṇ* 'sister'.
172. *báy* (I) 'today'.
173. *bãe(n)*^{1} (V) 'to spread leaves on a roof; thatch'.
174. *bã́en*^{2} (Adj) 'bad', cf. K. *baen* 'good'. Such occasional instances of meaning change are used by Korkus to claim that Nihals do not speak properly and are inferior people.
?175. *bãy-piyā* (NC) 'female papiha bird'. Its NH dialect name is *piyu*, *bāī* M. 'woman'.
L176. *beččerko* (V) 'to advise; thought', cf. H. *bicār*, M. *vičār*.
177. *beddiso* (NC) 'opaque'. This is *be-diso*. Reduplication of *d* by phonological conditioning. *diso* > cf. HM *dis* 'to appear, see'.
178. *bé* (V) 'to give'. (GRG *bɛḍ-*.)
179. *bédi* (N) dialect variant of 180.
180. *bédo*^{1} (N) 'enmity; evil'.
181. *bedo*^{2} (N) 'stick (wooden)'.
182. *bekki* (V) 'to sow; to reap', see 202.
L183. *bela* (V) 'to mix (dough, etc.)', cf. H. *belnā* 'to roll out bread'.
L184. *beldār* (N) 'mason', cf. M. *beldār* 'stone cutter'.
185. *bendi* (N) 'jungle', ? cf. *bir* SM.
186. *bénge áji* (NC) 'husband's elder sister'.
?187. *benga* (N) 'unit of length between thumb and the small finger with hand stretched', cf. M. *vengh*.
188. *be-nitto* (NC) 'blunt; not sharp'.
189. *bepa* (V) 'to thicken (gravy, etc.)'.
?190. *béra* (N) 'near relatives', cf. Md. *beḍa*.
L191. *berči* (N) 'spear', cf. M. *barči*.
192. *beríy* (V) 'to cut wood'; *-y-* causal.
193. *béri*^{1} (V) 'to gather'.
?194. *berí*^{2} (V) 'to beat a drum', cf. M. *bheri* 'a kind of a drum'; cf. SM *bad-*, *baq-*.
195. *béri*^{3} (V) 'to use', e.g. *dakhryaki bérika* 'uses left hand', cf. Ga. *ba()-*.

196. *bérko* (N) 'frog' in one dialect, 'tomcat' in another'.
L197. *berya* (Adj) 'naked', cf. K. *beṛia* 'mad'.
L198. *beryay-re* (NC) 'madman'.
199. *bérṭo* (N) 'husband's brother', cf. HM *bir*, *vir*.
200. *beso* (V) 'to whistle', cf. GR *sɨq*.
L201. *besuddi* (NV) 'faint, swoon', cf. M. *bešuddhi*.
202. *bétki* (V) 'to sow'; see 182.
203. *beṭṭo* (V) 'to die'.
204. *bétto* (V) 'to be destroyed by someone', e.g. *yedigoṭa béttoka davɛyna* 'bees killed by medicine'.
205. *béthe* (I) negative particle; *béʔ*, *béthela*, *béthel* are other forms given by the same informant.
206. *bethei* (I) 'not', see 54a; no other recorded occurrence recorded separately; *baiṭhna* (used as a euphemism for menstruation seclusion).
207. *béthel(a)* (I) see 205.
L208. *bether* (PP) 'underneath', cf. H. *bhitar*.
209. *bí* (V) 'to rise, stand, wake up', cf. K. *bid-*.
210. *bibol* (N) 'marking nut', *bibbā*.
211. *bicala páto* (VC) 'to be asked for in marriage'. (Among Nihals, the man invariably proposes.)
212. *bicá* (V) 'to ask a question'; see 176.
L213. *bida* (V) 'to see off; say goodbye', cf. H. *bidā*.
214. *biḍiko* (num) 'one'. Also found as *beḍe* (dialect variant), *baḍa*, *biḍik*, *biḍi*; cf. SG *boi*, GR *mui* (stem *miḍ-*).
L215. *bigar* (I) 'without', cf. HM *bigar*.
?216. *bigri* (N) 'name of a cooking vessel' (possibly a loan).
217. *bigi* (V) 'to blow (on fire); to blow (flute, etc.)', cf. SM *ped-* 'to blow a musical instrument'.
L218. *bigiri* (N) 'bangles', cf. M. *bangḍi*.
L219. *bijali batti* (NC) 'petromax lamp', cf. H. *batti* 'light', *bijli* 'electric'.
220. *biji* negative particle used only with imperative verbs.
221. *bijjok* (VC) 'to lie in wait for prey'. The first element is **bid-*, the second is obscure.
222. *bíkama* (V) 'to turn someone away', cf. K. *bae*.
223. *bikí* (V) 'to make'.
224. *bikiko* (V) 'to be made'.
L225. *bira* (N) 'betel leaf', cf. HMd. *biḍa*.
L226. *biri* (N) 'Indian cigarette', cf. HM *biṛi*.
227. *birṭo* (N) 'father-in-law'; see 199.
228. *birṭom* (N) 'father's elder brother; wife's elder sister; father-in-law'; see 199 and 227.
229. *biṭini* (NV) 'bed; to spread bedclothes', cf. NM *bil-*, *bel-* 'to spread', GR *bed-saq*.
230. *biṭil* (N) 'sand', cf. GR *bitiq* 'salt'.
231. *bithil* (N) dialect variant of 229.
232. *biṭṭhāwi* obscure verbal form; no recorded parallel: 'union, horizon, meeting', e.g. *agāso biṭṭhāwi khārā* 'sky and earth meeting on horizon'.
233. *biyā* (N) 'village'.
a. *biyā raywāsi* (NC) 'village dweller'.
L *raywāsi* loan, H. *rahivasi*.
234. *bogon* (NV) 'cry of a monkey', cf. SM *bor-*.
235. *bokki* (V) 'to construct, to tie something; bind', cf. SM *bod-*. *-ki-* in Nihali is added to root morphemes of verbs — its function is obscure. In Korku, *ki* is an intensifier.
236. *bokhārā* (N) 'loins'.
237. *bologo čikin* (N) 'caterpillar', cf. K. *cikhini* 'mosquito'. In Nihali, *bologo* does not occur by itself.
238. *bolor* (N) 'an edible root'.
L239. *bombil* (N) 'a variety of fish known in Indian English as Bombay duck, cf. M. *bombil*.
L?240. *bombo* (NV) 'cry of anguish at childbirth', e.g. *bombo biji* 'do not shout', cf. M. *bomb marṇe* 'cry loudly'. But M. uses it generally, not just for childbirth. Cf. R. *bombo* 'sexual intercourse'. Is it a loan into Marathi? For in M. it also describes ritual shout at Holi celebration.
241. *bo(m)moki* (N) 'siblings'; see 266 *bum(b)luy-*.
242. *bonde* (N) 'bud', cf. M. *bonḍ* 'bud on a cotton tree'.
243. *bondo* (V) 'separate grain from chaff'.
244. *bor* (Adj - V) 'hard', cf. M. *boboṛ*.
245. *boṛa* (N) 'scrotum', cf. K. *boḍa* 'penis'.
246. *borsal* (N) 'a species of bird'.
L247. *boṭo* (N) 'finger', cf. M. *boṭ*, K. *booṭo*.
L248. *botor* (N) 'hare; rabbit', cf. Hd.
249. *boṭhya* (N) 'a variety of fish'.
L250. *bóy* (N) 'grass, reed, weed, etc.', cf. K. *boe*.

a. *bóy ṭekḍāni* (NC) 'grazing ground'. Cattle are always taken by them on the hills to feed. Cf. M. *ṭekḍi* 'hill'.
251. *búči* (N) 'nape of the neck'.
L252. *buču* (N) 'breast, nipple', cf. K. *bucu*, Md. *buca*.
L253. *buḍi* (N) 'buttock'. Md. *buḍ* meaning same, also meaning 'bottom'.
L254. *búḍi* (V) 'to set (of the sun); to drown', cf. M. *buḍne* 'sink, drown, set'.
255. *buddi* (V) dialect variant of 254.
L256. *buḍḍo-kāmāy* (V) 'to dip into'.
257. *bukki* (V) 'to kiss'.
258. *buḍu* (V) dialect variant of 254, 255.
259. *buḍuluij* (N) 'bird's nest'.
260. *bulu* (N) 'bubble', cf. SM. *bubu-laɲ*, M. *buḍbuḍa*, H. *bulbula*.
261. *buluruij* (V) 'slip (handhold, foothold, etc.)'.
262. *budulwi* (N) dialect variant of 259.
263. *buluruy(e)* (V) dialect variant of 261.
264. *bulurik* (V) 'to be tickled'.
265. *bumli* (N) 'navel'.
L266. *bum(b)luy* (N) 'navel', cf. K. *bũbluj*, M. *bembi*. Cf. K. *bũbukɨ* 'siblings from the same navel'?
L267. *bundi*^{1} (N) 'tree trunk', cf. M. *bundha*.
268. *bundi*^{2} (N) 'nearby place'. This could be the same morpheme as 267.
269. *burāco* (Adj - V) 'to be spoilt', cf. H. *burā* 'bad, spoilt'.
270. *buṭu* (N) 'a kind of grass used for thatching'.

bh

L271. *bhāgā* (Adj) 'big, much'. Sk. *bhāgā* is a 'section, portion of' and is definitely not related.
   a. *bhāgā aba* (NC) 'grandfather'.
   b. *bhāgā biya* (NC) 'town', lit. 'big village'.
   c. *bhāgā dāi* (NC) 'elder brother', cf. Sk. *dāyabhāga* 'inheritance portion of the brothers'.
   d. *bhāgā dinā mānšo* (NC) 'old man'.
   e. *bhāgā dinā bāi* (NC) 'old woman'.
   f. *bhāgā dinā kol* (NC) 'old woman'.
   g. *bhāgā māy* (NC) 'grandmother'.
272. *bhāgāṭa* (Npl) 'ancestors' (never in sg.), cf. M. *bhagat* 'religious priest in the folk tradition who is possessed by "ancestors".'
273. *bhagān* obscure formation from 271: 'loudly'.
274. *bhāgeli* (Adj) 'long'.
275. *bhakkháru* (NC) 'a large group'. *bhag(a)* + *kháru*, lit. 'big herd of cattle', e.g., *bhakkáru saduṭa* 'assembly of medicine-men'.
276. *bhāgā pasārā* (NC) 'broad', also 'wide cervix'.
277. *bhāgel* see 274.
278. *bhagato* (N) '(white) magician'; see 272.
L279. *bhaggo* (N) 'stork', cf. M. *bagḷa*, Sk. *baka*.
?280. *bhajá* (V) 'to engage in activity'.
L281. *bhaji* (N) 'vegetables', cf. M. *bhaji*.
L282. *bhakkam* (Adj) 'solid', cf. M. *bhakkam*.
L283. *bhālā* (N) 'spear', M. *bhālā*.
284. *bhāl-kol* (NC) 'first wife (when alive)'. Now bigamy is impossible, but folk stories have polygamy.
   a. *bhāl kolnā palšo* (NC) 'first wife's children'.
285. *bham(b)(a)ra* (N) 'cough'.
L286. *bhānḍa* (N) 'pot', Md. *bhānḍa*.
L287. *bhānḍa-kunḍā* (NC) 'pots and pans', cf. M. *bhānḍi kunḍi*.
288. *bharan(g)wa* (N) 'earthworm'.
289. *bháte* (I) 'now'.
   a. *bháte menḍiṭi* (I) 'approaching night'.
290. *bhā́wri* (N) 'back'. It could be *bhaw-ri* (*re*, *-ri* dialect variation of personalizers).
   a. *bayi-bhā́wri* (NC) 'women folk', lit., 'backs of women' (cf. 'heads of cattle, etc.' in English).
291. *bhed(a)rā* (N) 'potato'.
292. *bheki* (V) 'to break'.
293. *bhek* (V) dialect variant of 292.
294. *bheltāni* (Adj) 'various', cf. M. *bhalta* 'odd', Md. *bhalta* 'various'.
L295. *bhendye* (N) 'okra (ladies fingers)', cf. Md. *bhende*.
296. *bhengi* (N) 'member of the sweeper caste of Hindus', cf. M. *bhangi*.
297. *bhér* (V) 'harrow'.
L298. *bhéra* (V) 'to fill', cf. M. *bharṇe*.
299. *bherjoli* (N) 'a small red jungle flower used as a spice in cooking'.
300. *bheriya* (V) variant of 298.
301. *bheriyākāmā* (V) 'to cause to fill, i.e., the pot fills; I fill the pot'.

302. *bherya* (Adj) ‘full’.
L?303. *bhetiri* (V) ‘to cut, to chop wood’, cf. Sk. *bhid* ‘cut’.
L304. *bhikko* (V) ‘beg’, cf. M. *bhik* (noun), *bhikkhu* Prakrit for Buddhist mendicant. Buddhist monks were contemptuously referred to as beggars later after Buddhism collapsed in India.
L305. *bhillā* (N) ‘kite (bird)’, cf. K. *bhilla*?
L306. *bhil* (N) ‘man from the bhil tribe’, cf. M. *bhilla*, H. *bhil*.
L307. *bhĩwárjun* (N) ‘names of two gods who live just outside any village under two trees’, cf. M. *bhimarjun* ‘two mythological characters from the Mahabharata’. There are other MBH characters occurring in Nihal folk-tales, e.g., *Nala*.
L308. *bholā* (Adj) ‘innocent’, cf. H. *bholā*.
309. *bhotāri* (N) ‘grass used for fodder’.
310. *bhowri* (N) ‘a conch-shell’.
311. *bhũa* (N) ‘trail of a wild animal’, cf. *pudaq* SM.
312. *bhuga* (N) ‘glass beads and/or pendant in a necklace’.
L313. *bhúki* (V) ‘bark (of a dog)’, cf. Md. *bhukṇe*.
L314. *bhúl* (V) ‘mistake, error; to forget’, cf. H. *bhūl*.
?315. *bhúm* (V) ‘to move about, to advise’, cf. H. *ghum-*.
L316. *bhuto* (N) ‘male ghost’, cf. MH. *bhūt*, K. *bhuuto*.
    a. *bhuto gathi jerei* (VC) ‘to be possessed by a ghost’.

c

L317. *ča* (N) ‘tea’, cf. Md. *ča*.
L318. *ča-petti* (NC) ‘tea-leaves’, cf. Md. *capatti*.
?319. *čačakāmā* (V) ‘to heat’, cf. Md. *catka*, see *cata* later.
320. *čačaw* (N) ‘grandson/daughter’, cf. in another dialect ‘wife’s brother’s son’.
321. *cačeṭo* (V) ‘to be wet’.
322. *cačom* (V) ‘to taste good’, e.g., *raččho cačomka* ‘juice tastes good’.
?323. *cacu* (N) ‘cock’s comb’. Sk. *cañcu*, M. *coc* is ‘beak’.
?324. *cačuko* (Adj - V) ‘hot’.
325. *čaggi* (V) ‘to aim, to shoot (an arrow)’.
326. *cagi* (V) dialect variant of 325, cf. M. *tɨɲ*.
327. *čágo* (N) ‘stone’.
328. *cahār* (NV) ‘echo, to hear an echo’.
329. *ćaka* (V) ‘to catch’, cf. SM. *cab-/jab-* ‘to catch’, see 335.
330. *caká* (V) ‘to climb’, see 335.
331. *čakāk* (V) ‘to be stationary’.
332. *cakāy-ajni* (VC) ‘to stoop, bend down’.
?333. *cakkān* (N) ‘whirlwind’.
334. *cakhāw* (V) ‘to sweep’, cf. SM. **juq-*, **joq-*.
335. *cākí* (V) ‘to cling’.
336. *cākini* (V) ‘to hear’.
337. *cāko* (V) ‘to climb’, see 330. Also in imperative form, one attested instance of *biji coka*.
338. *cālom* (V) ‘to beat the drum, etc.’, cf. GR. *luɲ-* ‘to thank/run’ ?
339. *caloŋ* (V) ‘thunder’, see 338 for cf.; e.g., *badra caloŋka* ‘clouds thunder’.
340. *calphẽj* (N) ‘covering of a cooking-vessel’, cf. SM. *dal*, K. *calpon* ‘palm of the hand’.
L341. *cambāli* (N) ‘small white fragrant flower similar to Jasmine’, cf. M. *cameli*, Md. *cambeli*.
342. *canā* (N) ‘dance’, cf. K. *su-sun* ‘dance’, SM. *susun* ‘to accompany’.
    a. *canā-dhol* (NC) ‘dance drum’, cf. M. *dhol* ‘drum’.
343. *cán* (N) ‘fish’, cf. K. *ka-*, K. *cade* ‘fish sp.’
344. *cāndā(w)* (N) ‘fish’, see 343, cf. K. *da?* ‘water’ *candata* (plural is more frequently used than singular).
345. *cāndāw(a)ti* (N) ‘neighborhood’, cf. 344 — ‘shares of fish’ ?
L346. *cāndi* (N) ‘silver’, cf. M. *candi*.
347. *capaṭ(a)* (V) ‘to chew’.
348. *cā-peṭh* see *ca* above.
349. *cāpiniku* (N pl) Nihal term for Korkus (derogatory), cf. *sani* ‘smell’.
350. *cāpir* (N) ‘an edible root’, cf. SM. *sa-* ‘classifier’.
351. *caraĩ* (N) ‘small river’.
352. *caré(k)* (V) ‘to cut’, cf. K. *cereq-*.
353. *carkhaḍ* (N) ‘waist’, cf. SM. *-sab* in *ki-sab*.
354. *cārko* (N) ‘black-faced monkeys’, cf. SM. *saq-*, *saraq-* ‘monkey’.
355. *cārmāru* (N) ‘centipede’, cf. SM. *-mar* ‘centipede, scorpion, stinging insect’.

356. *cārmāru* (N) 'yellow wild flower of many petals', cf. 355. It could be the same morpheme.
L357. *caṭā* (Adj - V) 'hot', cf. M. *catka*.
L358. *caṭāko* (N) '1/16 of a seer' (old measure before metric system was adopted in India). A seer is a little less than a kilogram. Cf. M. *čhatāk*.
359. *cáṭay* (V) 'to be burnt', cf. 357.
M360. *caṭni* (N) 'sauce', cf. MH. *catni*.
361. *caṭo* (I) 'quickly', cf. M. *catkan*, H. *cat-se*.
362. *cáto* (NV) 'hunger; to be hungry'.
363. *caṭṭā* (Adj - V) 'hot', see 357.
L364. *caturāy* (NV) 'cleverness', cf. M. *caturai*.
365. *cāwak* (V) 'to be afraid'.
366. *cāwgó* (NV) 'fear'.
a. *cāwgo manšo* (NC) 'a coward'.
367. *cāwki* (N) 'heart of a lamb'.
368. *cauni* (N) 'temple'.
369. *cayni* (num) 'first'.
370. *ceki* (V) 'to hold, arrest'.
371. *cekí(k)* (V) 'to catch (a ball, etc.)'.
372. *cekki* (V) 'to pursue'.
373. *cekini* (V) 'to hear'.
374. *cekoṭo* (V) 'to chop meat', cf. GU *seb-* 'to slaughter'.
375. *cekṭo* (N) 'knife', cf. 374.
?376. *celkeye* (N) 'small market basket carried on head'.
377. *cenda* (V) 'to throw'.
378. *cended* (N) 'moon'.
379. *céyni* (num) 'first, previously', see 369.
380. *cerki* (V) 'drop a thing'.
381. *cerke* (V) 'to fall'.
?382. *cera* (V) 'to cut (with scissors, etc.)', cf. K. *cereq*, M. *cirṇe*.
383. *cerci* (V) 'to sprinkle'.
L384. *ceré* (V) 'to cut (hair)', cf. K. *cereq*.
385. *cere-mere* (VC) 'to fall' echo formation 'm' as echo in Md.
386. *cergó* (V) 'to run', cf. K. *saṛub-*.
387. *céṭer* (N) 'folded hands (as in prayer)'.
388. *ceṭerbako* (VC) 'to give birth'.
389. *cether* (N) 'a poisonous lizard'.
L390. *cicca* (N) 'the tamarind fruit', cf. Md. *cicca*.
a. *cicaḍḍo* (NC) 'tamarind tree'.
391. *ciciḍoke* (N) 'a species of birds', cf. K. *tid-* 'bird'.
392. *ciciler* (NC) 'a tree used for medicinal purposes'.
393. *cigām(a)* (N) 'ear/s' int. plural.
L394. *cikal* (N) 'mud', cf. Md. *cikal*.
395. *cikin* (N) 'worm, small winged insect', cf. Munda **si/se* 'louse'.
396. *cikinā* (N) 'threadworms', pl. *cikinṭa*, see 395.
L397. *ciko* (N) 'sticky substance; sap of a tree', cf. M. *cik*.
L398. *cikaṭo* (NV) 'to be selfish, miserly', cf. M. *cikaṭ* 'miser'.
399. *cilar-ṭa* (N pl) 'lice' always used in plural, also used for body lice; cf. Munda **si-*, **sen-*.
400. *cilatak* (V) 'to cut up in pieces', cf. M. *ciraḍṇe* 'crush'.
401. *cilaṭi*^{1} (N) 'a species of bush'.
402. *cilaṭi*^{2} (N) 'lower half of the grinding stone (the domestic grinder is made of two flat stones one on top of the other with a rotating axis, which moves the top half only).
403. *cile* (N) 'chicken', cf. Md. *cili*.
404. *cilir* (V) 'to slip (a foot)'.
405. *cilom* (N) 'tobacco pipe'.
406. *cimni-tel* (NC) 'kerosene', lit. 'oil for the wick lamp'.
407. *cimmi* (N) 'kerosene oil lamp', cf. 406 *cimṇe*.
408. *cíp* see 408a.
a. *cíp kogo* (NC) 'cobra'.
409. *cippo* (V) 'stay, stand'.
410. *cirempār* (I) 'evening time', cf. *pār* 'time' in M. > *prahan* Sk.
411. *ciṭi* (N) 'goddess who visits new-born babies', cf. MH. *sati*.
412. *ciṭipara* (NC) 'parting of the hair', cf. 411.
L413. *ciṭṭhi* (N) 'letter', cf. Md., H. *ciṭṭhi*.
414. *ciṭo* (N) 'engagement for marriage'.
415. *cĩya* (N) 'seed of a tamarind fruit'.
416. *có* (NV) 'urinate'.
417. *co(g)gom* (N) 'pig'.
418. *cojona* (N) 'nose'.
419. *cokob* (N) 'leaf' also 'a clan name'.
420. *cole* (N) 'cooking stove using firewood', cf. M. *cul*.
?422. *comṭa* (N pl) Korkus call Nihals by this term. It means 'thieves' (cf. M. *cor-*). It is using *-ṭa* Nihali plural suffix, and Nihals know the term; hence, it is included here.

423. *cón* (N) 'nose', cf. 418, but also 'mouth' for some informants; 'mouth' in the dialect I use as above.

424. *cón-dhanḍi* (NC) 'bridge of the nose'; second element may be a loan, cf. M. *danda* 'stick'.

425. *cón-popa* (NC) 'nostril'.

426. *cópo* (N) 'salt'.

427. *copṭi*^{1} (N) 'a drop of water'.

L428. *copṭi*^{2} (V) 'to live together', cf. M. *sobat.*

L429. *cor* (N) 'thief', cf. MH. *cor.*

430. *cokoṭo* (N) 'axe'.

L431. *corgi* (N) 'a woman's top garment', cf. MH. *coli.*

432. *corṭo* (NV) 'blood; bleed'.

L433. *cosa-ṭiŋ* (VC) 'to think', cf. K. *soca-tiŋ*, H. *socna.*

434. *cucórmug* (NC) 'house lizard'.

435. *cucu* (N) 'drinking vessel', cf. M. *coc* 'beak'.

L436. *cunā* (N) 'lime', cf. H. *čunā.*

437. *curi* (N) 'name of a fish'.

438. *cutiñ* (N)

439. *cuṭṭi* (V) 'to pound again'.

440. *cuṭuli* (N) 'musk rat' (considered auspicious). Name of a clan.

ch (see 1520-1552)

d

L441. *daba-ṭiŋ* (V) 'to press', cf. K. *daba-ṭiŋ*, H. *dabna.*

L442. *dabba* (N) 'storage receptacle with a lid', cf. M. *dabā.*

L443. *dabo-dabo* (N) 'halting speech', cf. H. *dabana* 'to check, halt'.

L444. *ḍabra kheto* (NC) 'field on level ground', Md. *dabar*, H. *khet.*

L445. *dada* (N) 'mother's brother; father', cf. M. *dada* 'elder brother'.

L446. *daḍi* (N) 'beard', cf. M. *dadhi.*

447. *daddhi* (N) 'chin', cf. 446.

?448. *dai* (N) 'brother; husband of husband's sister', cf. Sk. *dayad.*

449. *daka* (I) 'after that'.

L450. *dakara* (NV) 'burp', cf. H. *dakar.*

L451. *dawāy* (N) 'medicine', cf. H. *davāi.*

L452. *dakā́yn* (V) 'to show' causal of 'to see' *dakan*, cf. H. *dekhna*; expected form is **dakānāy* for causal.

L453. *dākhrya* (N) 'left-handed person', cf. M. *dāvkhorā*, Md. *dakhurā.*

L454. *dakkāre* (NV) 'to hiccough', cf. H. *dakār*, see 450.

L455. *dakāri* (N) 'old in age (but not feeble)', cf. RG. *dukri* 'old', cf. Gujarati *dokra*, *dokri* same meaning.

L456. *dāko* (N) 'a small drum', cf. Md. *dakkā*, M. *dankā.*

L457. *dam* (N) 'breath', cf. H. *dəm.*

L458. *dām* (N) 'money', cf. H. *dam.*

459. *dā́mborā* (NC) 'a species of tree'.

460. *dān(a)* (V) 'was; were' (forms of 'to be' are irregular).

461. *dāná* (N) 'clitoris'.

462. *dānḍi* (N) 'straight stick', cf. Sk. *danḍa.*
  a. *ubā dānḍi pākṭo* 'spine'.
  b. *pākhoroko danḍi* 'bone of thigh, bone of arm'.

L463. *dānḍi punam* (NC) 'full moon day in the month of phalguna; main day of the Holi festival' (Holi is known for sexual frolicking), cf. M. (vulgar) *dānḍa* 'phallus', Md. *punam* 'full moon day'.

L464. *dānḍo* (N) 'handle'.

L465. *ḍāngi* (N) 'branch', cf. Md. *daŋg.*

L466. *dāngārā* (N) 'cucumber', cf. Md. *dangar* 'gourd'.

467. *ḍāngo* (N) 'branch', cf. 465.

?468. *darā́* (NV) 'call of a bird', e.g., *koyal darā́keḍini* 'cuckoo is singing', cf. M. *ḍarkāḷi* 'loud noise'.

L469. *dāri* (N) 'gram', cf. Md. *dāri*, M. *dāl.*

?470. *darom* (V) 'to escort a person (to show respect)', K. *darom* 'outside courtyard', cf. M. *ḍar* 'fear'.

L471. *daski* (N) 'birth; delivery', cf. H. *das* 'ten'; ritual impurity of ten days connected with birth.

L472. *dāṭi* (Adj) 'thick', cf. M. *dāṭi.*

L473. *dawā* (N) 'medicine', cf. H. *dava.*

L474. *dāwrā* (NV) 'harrow', cf. Md. *davrā* 'wooden plough'.

475. *dáy* (N) 'way; road'.

476. *-dāyā* (pp) 'for', e.g., *lá daya kóyi* 'brought for you'.

477. *deḍḍa* (N) 'frog', cf. K. *dedda?*, SM. *-dɛd*, *-dɛb*.

478. *dej bābārā* (NC) 'an edible root'.

479. *delen* (V) 'to drink'.

480. *dene* (pronoun) 'his', cf. SM. gen. *-nd*, K. *di-*.

481. *déngara* (N) 'a huge log of wood'.

482. *deŋgni* (N) 'echo'.

L483. *dewtā* (N) 'sun', M. *devatā*, *dev* used to refer to the sun.

484. *ḍibi* (N) 'a poisonous root'.

L485. *dil jeréi* (V) 'to be in love with', cf. H. *dil* 'heart', Marathi *jaḍne* (V).

L486. *din* (N) 'day', cf. H. *din*.

487. *ḍimbo* (N) 'paw of an animal.

L488. *diŋkāre* (NC) 'gum tree', cf. M. *diŋka* 'gum'.

L?489. *disoka* (pronoun) 'another; transparent'.

490. *dísu* (V) 'to appear'.

L491. *diwāli puja* (NC) 'Diwali festival'.

492. *diwi* (V) 'to wave (a hand, etc.)'.

L493. *diyā* (NV) 'lamp; light; to be late'.

L494. *ḍo* (I) 'and; then', K. *ḍo*.

495. *ḍobā* (N) 'bull'.

L496. *ḍobhar* (N) 'a puddle of water', cf. Md. *ḍabrā*.

L497. *ḍohor* (N) 'the chamars (a Hindu caste traditionally shoemakers)', cf. M. *dhor* 'cattle' (also used as a caster name).

498. *dokčó* (V) 'to bleed'.

499. *ḍoko* (V) 'to hiccough'.

500. *ḍokom* (V) 'to wear (clothes)'.

501. *dokso*^{1} (N) 'flood'.

L502. *dokso*^{2} (N) 'caution', cf. M. *dakša*.

503. *dokuma* (N) 'a leaf cup'.

L504. *ḍoktar* (N) 'a doctor (practicing allopathy)'.

L505. *dole* (N) 'cardamom', cf. M. *veldode*.

L506. *ḍongā* (N) 'a red ant', cf. M. *dongḷa*.

?507. *dongor* (V) 'to rest'.

508. *ḍongor bérko* (N) 'wild cat'.

L509. *dorā* (N) 'string', cf. M. *dorā*.

510. *ḍotāko* (N) 'edible root'.

L511. *dotkā* (N) 'a kind of gourd', cf. M. *doḍka*.

512. *ḍubumbí* (NC) 'abdomen'.

513. *ḍuburu* (N) 'a bamboo basket tied to the waist when catching fish, for collecting them'.

L514. *dudo* (N) 'milk', cf. M. *dud*, K. *didom*.

515. *ḍugi* (N) 'red-faced monkey', cf. K. *dugi*.

L516. *dukāni* (N) 'bed-bug', cf. M. *dhekuṇ*.

L517. *ḍukrimāy* (NC) 'grandmother', also 'father's older sister'; it is possible to marry older sister's daughter in South India in many castes.

518. *dumā* (N) 'lizard'.

519. *dupu* (N) 'marigold', cf. K. *dhupu*.

520. *duri* (V) 'to flow'.

L521. *dusara* (num) 'second', cf. M. *dusra*.

L522. *dusman* (N) 'enemy', cf. H. *dušman*.

523. *dusorā murat* (NC) 'different thing', cf. 521.

524. *dusra sāl* (NC) 'year before last', cf. 521, M. *sal* 'year'.

525. *ḍuṭu* (N) 'a bamboo basket', cf. Md. *duṭi*, *duḍi*.

dh

526. *dhā* (I) 'there'.

527. *dhāblā* (N) 'name of a bird', cf. K. *dabili*.

L528. *dhanāko* (N) 'bow', cf. M. *dhanušya*.

529. *dhāpri* (N) 'bank of a river; clan name'.

L530. *dhār* (N) 'sharp edge', cf. M. *dhār*.

L531. *dharom manšo* (NC) 'a generous man'.

L532. *dhaṭṭo* (NV) 'brave; to dare', cf. Md. *dhaṭṭa*.

533. *dhauni* (V) 'to cure (of a sickness)'.

534. *dhāvā* (Adj) 'distant'.

535. *dhāwči* (N) 'name of a bird'.

536. *dhāwrā* (V) 'gum tree'.

537. *ḍheḍhne* (N) 'bumble bee'.

L538. *dhẽya* (N) 'curds', cf. Md. *dhẽi*.

539. *ḍhengi* (Adj) 'erect; stand straight'.

L540. *ḍhepi* (N) 'wall', cf. Md. *ḍhepla* 'clod'.

541. *dherti ékhāy* (NV) 'earthquake', cf. H. *dharti* 'earth'.

L542. *ḍhilā* (Adj - V) 'loose', cf. M. *ḍhila*.

L543. *dhobi* (N) 'washerman', cf. HM. *dhobi*.

544. *ḍhoḍhór* (N) 'skeleton', cf. M. *dhor* 'cattle'; partially reduplicated form *ḍho-ḍhór*.

L545. *dhokā* (NV) 'danger; betray'.

546. *dhokanān* (N) 'large mosquito'.

L547. *ḍhol* (N) 'drum', cf. M. *ḍhol*.

L548. *dhollā* (Adj - V) 'hollow', cf. M. *ḍholi* 'hollow of a tree'.
L549. *dholkya* (N) 'drummer', cf. M. *dholkya*.
?550. *dhonḍum* (N) 'mouth of an animal', cf. M. *tonḍ*.
551. *dhongāri* (N) 'a type of grass'.
L552. *dhoporā* (N) 'buttocks', cf. M. *dhopar* 'knee', but K. *dhopar* 'buttock'.
L553. *dhor* (N) 'cattle', cf. M. *dhor*.
L554. *dhot carāvre* (NC) 'one who tends cattle' (*re-* is a personalizer).
L555. *dhow* (NV) 'wash; bathe', cf. Md. *dhau*.
L556. *dhũār* (N) 'frost, mist', cf. H. *dhũa*.
557. *dhuḍi* (N) 'a variety of fish'.
L558. *dhuk-dhuk* (NV) 'beat of a heart', cf. M. *dhuk-dhuk* 'beating of heart'.
L559. *dhullā* (N) 'dust', cf. Md. *dhulla*.
L?560. *dhundi worā* (N) 'dust storm'.
561. *dhupkā* (N) 'collision'.
L562. *dhurā* (N) 'smoke', cf. M. *dhur*.
563. *dhúrā* (N) 'youngest person in the group'.
L564. *dhũyā* (N) 'smoke', cf. H. *dhũa*.

e

N.B.: All forms beginning with *e* optionally have palatal onset. In the speech of any one speaker, the forms with [ye] and [e] occur in free variation; but almost everyone uses some forms in [e] beginning.

565. *e* (Adj) demonstrative third person.
566. *edé* (N) 'pain'.
567. *éden* (V) 'to hurt, feel pain'.
568. *édugo* (N) 'house-fly'.
569. *eger* (V) 'to remove; to shave'.
570. *ejer* (N) 'boy; son'.
571. *ekhāy* (V) 'to shake in the breeze'. Also see *dherti ékhāy* 'earthquake'.
572. *ekhelkā* (pronoun) 'alone; by oneself'.
573. *ekher* (V) 'to be spilled'.
574. *elā-elā* (I) 'quickly'.
575. *én* (V) 'to come, approach'.
?576. *engan* (N) 'egg-plant (brinjal)', cf. H. *baiŋgan*.
L577. *enger* (N) 'burning coals; flame', cf. M. *angar* 'fire'.
578. *engā* (pronoun) 'that one'.
579. *engerkāri* (N) 'matchstick'.
580. *épṭa* (N) 'bees' (always plural).
581. *eptāvār* (NC) 'bee-hive'.
582. *ér* (V) 'to go'.
583. *ere* (pronoun) 'this one', K. *i(n)- ~ e(n)-*.
584. *erṭā* (V) 'to thrash (the grain)'.
585. *etān* (pronoun) 'this'.
586. *eté* (pronoun) 'he'.
587. *etlá* (pronoun) 'they'.
588. *etre* (pronoun) 'you'.

g

L589. *gādáw* (VN) 'to bury; cemetery', cf. 590.
L590. *gādāy* (NV) 'to bury', cf. M. *gadṇe*.
L591. *gadri* (N) 'donkey', cf. M. *gadhaḍa*.
?592. *gāghāy* (N) 'scar', cf. M. *ghav*.
L593. *gājre* (N) 'carrot', cf. M. *gajar*.
L594. *gāl* (N) 'cheek', cf. M. *gal*, K. *gal*.
L595. *gālā* (N) 'trap', cf. M. *gaḷ*.
L596. *galāté* (V) 'to dissolve', cf. M. *gaḷne*.
L597. *galām* (N) 'braid', K. *galàm*.
598. *gánda* (V) 'to cut'.
L599. *gandā* (N) 'stupid', cf. H. *ganda* 'bad'.
600. *gāngrya* (N) 'a non-tribal'; also contemptuously used to refer to Nihals who now speak Korku.
601. *gar gatál* (N) 'testicles'.
L602. *garā*^{1} (N) 'bullock cart', cf. Md. *gaṛa*.
603. *garā*^{2} (N) 'inside of an egg-ball; testicles'.
L604. *gardan* (N) 'neck', cf. H. *gardan*.
L605. *gārgoṭi* (NC) 'hailstone', cf. *gār*.
606. *gaṭā́l* (N) 'penis'.
a. *gaṭā́l koṭṭo* (V) 'to castrate'.
L607. *gāṭhi* (V) 'to string'.
L608. *gayrā* (Adj - V) 'loud, deep', cf. H. *gahirā* 'deep'.
609. *geliyā* (Adj - V) 'unworthy; useless'.
610. *genjo* (I) 'enough'.
L611. *gele* (N) 'maize', cf. K. *gele* 'ear of corn (maize)'.
L612. *geri* (N) 'fishhook', cf. MH. *gaḷ*, *gal*.
613. *gethāy* (N dual) 'dog and bitch'.
614. *geyi* 'supporting beam of the roof'.
615. *gichā* (Adj - V) 'tiny'.
L616. *girbo* (N) 'poor or gentle man', cf. HM. *garib*.
L617. *girjā* (N) 'church', cf. Md. *girjaghar*.
L618. *giru* (N) 'red ochre', cf. M. *geru*.

L619. *giryān* (N) 'eclipse', cf. Md. *giran*.
620. *giṭā* (N) 'brother and sister; sibling, younger siblings of wife'.
621. *gobṭi* (N) 'bun of hair'.
622. *godāl*^{1} (Adj) 'round; top stone of a grinding stone'.
L623. *godāl*^{2} (N) 'cooking oil', cf. Md. *godel*.
L624. *goday* (NV) 'tattoo', cf. Md. *godne*.
?625. *gogāṭi* (N) 'snail', cf. M. *gogalgay*.
626. *gogoṭi* (N) 'a kind of a bird'.
L627. *gohũ* (N) 'wheat', cf. M. *gahu*.
?628. *gohum* (N) 'bride price', cf. 627.
L629. *gol* (N) 'beam'.
L630. *golá* (V) 'gather; rake leaves', cf. K. *gola-* 'to collect'.
631. *golakama* (V) 'heap'.
L632. *goli* (N) 'tablet', cf. Md. *goli*.
L633. *golga* (N) 'earwax', *golŋga*.
L634. *golor*^{1} (N) 'a slingshot', cf. M. *galor*, K. *golor*.
635. *golor*^{2} (Adj) 'curved', cf. *gol*.
636. *gol-pago* (N) 'stable for horses', cf. M. *paga*, *ghod-paga*?
637. *gor* (N) 'jaggery' (a kind of sugar).
L638. *gorha* (N) 'young he-calf', cf. M. *gorha*.
L639. *gorsi* (N) 'buttermilk', cf. M. *goras* 'milk'.
640. *goradi* (N) 'poison made from the bark of goradi tree'.
641. *gorya* (N) Korku-speaking Nihals use this term to refer to Nihali-speaking Nihals.
642. *gothi* (N) 'clan', cf. M. *gotha* 'cowpen', *gotul* 'clanhouse among M.P. tribals (Gonds)'.
643. *gotor* (N) 'pumice stone'.
L644. *goyi* (N) 'tablet'.
645. *gúdo* (N) 'marrow; brain', cf. H. *guda* 'pulp'.
646. *gugāri* (N) 'a male', *gogo* 'bird'.
647. *gugudo* (N) 'an edible root'.
L648. *gulāmb* (N) 'rose', cf. M. *gulab*.
649. *gulguṭi* (V) 'gurgle'.
650. *gulmor* (N) 'the Golden Mohur tree'.
651. *gumpli kamāy* (VC) 'to tickle'.
L652. *gunḍi* (N) 'a button', cf. M. *gundi*.
653. *gursā* (N) 'anterior tibial ridge', also 'bones of hands'.
654. *guṭṭu*^{1} (V) 'join'.
654a. *guṭú*^{2} (V) 'to thread'.

gh

L655. *ghām* (N) 'sunshine', cf. Md. *gham*.
L656. *ghāmo din* (NV) 'hot, dry season (summer)'.
L657. *ghāní* (V) 'to smell', cf. Sk. *ghraṇa*, cf. Md. *ghaṇṇa*.
658. *ghāniki* (Adj) 'fragrant', see 657.
659. *ghanimani* 'round and round'.
660. *gharāndi* (N) 'a flat round clay plate'.
661. *ghātā* (V) 'to arrive'.
662. *ghātān ér* (VC) 'to earn'.
?663. *ghatthā* (N) 'word', cf. M. *gatha*.
664. *ghedyāl* (N) 'watch'.
L665. *gheri gheri* (Adj) 'frequent', cf. M. *ghadi ghadi*.
L666. *ghẽu* (N) 'ghee'.
L667. *ghóna* (Adj - V) 'many', cf. R. *ghana*.
L668. *ghonātiŋ* (V) 'to do in great style'.
L669. *ghum* (V) 'move about', cf. H. *ghum-*.
L670. *ghunghru* (N) 'ankle bell'.
671. *ghúr* (NV) 'anus; to excrete'.
L672. *ghuṭāri* (N) 'deer', K. *ghotaṛi*.

h

Word final *h* and intervocalic *h* in a single morpheme is phonetically [ʔ].

L673. *ha* (pronoun) proximate demonstrative stem, cf. M. *ha* 'he (proximate)'.
L674. *hã* (I) 'yes', cf. H. *hã*.
675. *hã́go* (NV) 'laugh'.
L676. *hākko* (NV) 'call', cf. M. *hak*, K. *haako-*.
677. *hakko orṭāy* (VC) 'to echo', cf. M. *hak*, M. *ulta* 'reverse'.
L678. *hālhāl kām* (NC) 'torture'.
L679. *halkā pāwšo* (NC) 'dew'.
L680. *hān* (pronoun) 'that', see 673, K. *han ḍi-* 'that'.
L681. *hārātiŋ* (V) 'to defeat'.
682. *hardo* (N) 'turmeric'.
683. *hare-hare* (I) 'noise made to shoo cattle away'.
684. *harjo* (N) 'mushroom'.
685. *hāri* (N) 'a space between two planted rows which has been weeded'.
686. *harpo* (V) 'to wander'.

687. *hāru* (V) 'bite'.
688. *hāruy* (V) 'to be bitten'.
L689. *hāt* (N) 'hand', cf. M. *hat*, H. *hāth.*
690. *hāti* (V) 'to be incomplete'.
L691. *haṭo* (N) 'market', cf. H. *hāṭ.*
692. *hawā* (N) 'locust'.
693. *hāyom.*
694. *-hela* verbal suffix.
L695. *hellā* (N) 'he buffalo', cf. M. *hela.*
696. *helle-helle* (Adj) 'this much; a handful (of grain, etc.)'.
697. *heŋk* (V) 'to converse'.
698. *he-re* (pronoun) 'this one'.
699. *hikaṭ* (V) 'to spin as a top', cf. *ekhay.*
700. *hiltel* (pronoun) 'these two'.
701. *him* (N) 'cold'.
  a. *him din* (NC) 'winter'.
L702. *himoṭo* (Adj - V) 'courageous', cf. M. *himmat.*
703. *hindyān* (V) 'to aim'.
704. *hingā* 'this side; towards', cf. K. *(h)iŋgàn* (*[h]iŋga-èn*).
705. *hiṭini* 'here'.
706. *hiŋki* 'here'.
707. *hiyān* 'this much'.
708. *ho* (pronoun) 'he'.
709. *hoʔ* negative particle.
710. *holoy* (V) 'shake', cf. M. *halṇe*, K. *holoe.*
711. *hombo hombo* 'walking on all fours'.
712. *hombo* 'crooked', cf. K. *hõbàʔ-.*
L713. *honā*^{1} (V) 'to be' (only in certain constructions), cf. H. *hona*, e.g., *here eŋki bethé honā* 'I do not (want) him'.
honā^{2} (V) 'to accept'.
714. *hondār* (N) 'rat', cf. M. *undir* — considered to be an old loan into Skt.
L715. *horyā* (N) 'a bird', cf. M. *hola.*
716. *hoṭākā* 'there (away from speaker)'.
717. *hoṭi* 'far away'.
718. *hoṭikel* 'those two'.
719. *hoṭin* 'there'.
720. *hoṭiyān* 'that much'.
721. *hotiāʔ* 'that thing'.
722. *how* 'remote'.
723. *howeṭlā* 'these'.
724. *hũā* 'of this type'.
725. *húndā* (Adj - V) 'deep'.
726. *hunḍār* (V) 'cook, bake', cf. K. *hõḍar-.*
727. *hukuṭ* (V) 'to hiccough'.
728. *hur* (V) 'to gore; attack with horns'.
L729. *husārā* (Adj) 'experienced, clever'.
730. *huṭiki* 'there'.
731. *huyām* (V) 'praise; boast, gossip'.

i

y onset optionally.

732. *ičhā* (V) 'pinch'.
?733. *igin* (V) 'to believe, trust'.
734. *ikhaṭ* (V) 'to turn around oneself', see *hikaṭ* and *ekhāy.*
L735. *ilāyci* (N) 'cardamom', cf. H. *ilayci.*
736. *ilur* (N) 'husband's younger brother', cf. K. *ilur.*
737. *imbiri* (V) present tense of 'to be'.
738. *imni* (V) 'to be'.
739. *ingin* (pronoun) 'we' (plural).
L740. *inkāri* (N) 'weaver', Md. *iṇkar.*
741. *ír(i)* (num) 'two'.
  a. *ír janki bāro* 'duet'.
  b. *yir nārā* 'gullet'.
742. *irār* (num) 'two'.
743. *irā* (N) 'sickle', cf. K. *(h)ir.*
744. *irāy* (V) 'to cut'.
745. *irkār* (V) 'to roast'.
L746. *iskul* (N) 'school', cf. HM. *iskul.*
747. *iṭi* (PP) 'with regards to'.
748. *iṭieri* 'just now'.
749. *itkā* 'here'.
750. *iṭikel* (pronoun) 3rd person dual.
L751. *itawār* (N) 'Sunday', cf. Hd. *itwār.*
752. *iyér* (V) 'to come', see *eḍ.*
753. *iwā* 'like this'.
L754. *iwāy* (N) bride's father and groom's father are *iway* of each other, cf. Md. *iwai*, M. *vyahi.*

j

755. *-jā* (through the female line personalizing suffix), see 806.
L756. *jabrāy* (V) 'rape', cf. H. *jabri.*

L757. *jādā* (Adj) 'more', cf. H. *ǰada*.
L758. *jaḍi* (N) 'root', cf. M. *jadi*.
759. *jaḍe* (V) 'to join', cf. H. *jaḍna*.
L760. *jādu tembryā* (NC) 'were-tiger' (cf. 'werewolf').
L761. *jāduko* (N) 'magician'.
762. *jagali malā́y* (NC) 'spider'.
a. *jagali malā́ynāvār* (NC) 'web of a spider'.
L763. *jālā* (N) 'net', cf. Md. *jala*.
L764. *jaldi* 'quickly', cf. H. *jaldi*.
L765. *jalom* (N) 'birth', cf. Md. *jalam*.
766. *jalóm* (V) 'to thunder'.
767. *jālú* (V) 'to climb down (from a tree, etc.)'.
L768. *jā́lu* (NV) 'fever', cf. Md. *jaḷ*.
L769. *jambu* (N) 'blackberry', cf. M. *jambhul*; also a clan name (is it an old loan into Sk.?).
770. *jā́mo* (V) 'to measure'.
771. *jāmi* (N) 'buyer'.
L772. *janā* (N) 'person', cf. M. *jan* used with numerals, e.g., *yiṅ jana mānšo* 'two people'.
773. *jāngri* (N) 'foreigner; non-tribal'.
774. *jappi* (Adj) 'thick'.
L775. *jarā*^{1} 'just a little', cf. M. *jaralsa*.
L776. *jarā*^{2} (NV) 'fever', cf. M. *jvar*.
777. *jārān* (N) 'crab'.
L778. *jarmali bhāndā* (NC) 'cooking utensils made of German silver (aluminum alloy)'.
L779. *jātā* (N) 'husband's brother's wife'.
L780. *jathā* (N) see 779.
L781. *jāto* (N) 'caste', cf. MH. *jat*.
L782. *jatrā* (N) 'village fair', cf. M. *jatra*.
L783. *jawān* (N) 'young', cf. H. *jawan*.
a. *jawān worāṅ* (NC) 'eligible male'.
784. *jā̃w* (V) 'to choose'.
L785. *jekhām* (NV) 'wound', cf. M. *jakham*.
L786. *jeḷbi* (N) 'a sweet', cf. M. *jilbi*.
787. *jeré* (V) 'to become; to be ready' (auxiliary verb of importance).
788. *ǰiki(r)* (N) 'eye'.
789. *ǰiki bí* (V) 'to suffer from sore eyes'.
790. *jiki-kāpri* (NC) 'eyelid'.
791. *jiki ratā́y* (V) 'reddening of eyes'.
792. *jiki tā́pne* (NC) 'eye-brow' (cf. 1572).
793. *jiki yā̃to* (N) 'tears' (cf. 92).
794. *jilgutiñ* (N) 'earthworm', cf. K. *jilŋgod*.
795. *jiryāngā* (N) 'tomato'.
L796. *jiw* (N) 'soul'.
797. *jogo* (N) 'leech', cf. H. *jõk*.
798. *jogāgo* (N) 'gum', cf. *jer* (Korku).
799. *jochí* (V) 'to bake'.
800. *joppo* (N) 'water'.
801. *júdo* (N) 'beams of a roof'.
802. *judó* (V) 'to betroth'.
L803. *jummú* (V) 'to answer a question'.
804. *jó* (pronoun) 'I'.
L805. *jawāy* (N) 'son-in-law', cf. M. *jawai*, K. *jaway*.
806. *jawāy-ja* (NC) 'one's own son-in-law'.
L807. *jeher* (N) 'poison', cf. H. *jahar*.
L808. *jikiraphid* (V) 'blink; wink'.
L809. *jitta-jinna*, cf. M. *jirka-titka*.
L810. *jiwta* (Adj) 'alive', cf. M. *jivanta*.
811. *jiwtaddo* (N) 'living tree (opp. of dead)'.
812. *joda-jerei* (N) 'engaged couple'.
813. *jodu* (N) 'bamboo'.
814. *joppo-ikhat* (NC) 'whirlpool'.
815. *junu* (NV) 'broom; sweep', cf. K. *j-un-u* 'broom', *jug-* 'to sweep'.

jh

L816. *jhagarā* (N) 'quarrel', cf. H. *jhagḍa*.
?817. *jhakā-jhokú* (V) 'to stagger'.
L818. *jhakkān* (N) 'lid', cf. M. *jhakaṇ*.
819. *jhāllyā* (N) 'peacock', cf. Gu. *jhalia* (male).
L820. *jhanḍā* (N) 'flag', cf. H. *jhanda*.
?821. *jhāpnā́* (V) 'to cover', cf. M. *jhakṇe*.
822. *jhéli* (V) 'to defend against attack'.
823. *jher* (V) 'to swing'.
824. *jhiŋgā* 'there'.
L825. *jhirā* (N) 'spring, brook', cf. M. *jhara*.
L826. *jholā* (N) 'swing', cf. H. *jhalna*.
?827. *jhuri* (V) 'to rock a cradle', cf. H. *jhulna*.
?828. *jhoryānṭa* (N pl) 'cockroaches', cf. M. *jhuraḷ*.

k

?829. *kā* (N pp) 'towards', cf. H. *ko*.
L830. *kacarā* (N) 'garbage', cf. M. *kacra*.
L831. *kāḍakko* (Adj) 'strong, concentrated', cf. M. *kaḍak*.
832. *kājar* (N) 'top'.
833. *kajiṭo* (VN) 'to steal; thief'.

834. *kā́go* (V) 'to taste bitter'.
835. *kāggo* (N) 'mouth'.
836. *kāgo-kāgo jhagra* (N) 'quarrel'.
L837. *kākā* (N) 'uncle', cf. H. *kākā*, also M. *kākā*.
L838. *kāki* (N) 'aunt', cf. H. *kākī*, also M. *kāki*.
L839. *kākato* (N) 'paper', cf. M. *kagad*, H. *kagaj*.
840. *kākhén* (VN) 'comb'.
L841. *kākri* (N) 'cucumber; name of a clan'.
842. *kākhā* (V) 'to carry in arms'.
843. *kakheŋki* (N) 'comb', see 840.
L844. *kal* (N) 'noise'.
855. *kalkal* (N) 'noise'.
?856. *kalāṭ* (N) 'wife'.
856. *kalāṭṭel* (N) 'brother-in-law and sister-in-law'.
857. *kalāṭṭēyā* (N) 'wife's brother'.
858. *kalṭo* (N) Nihali name for themselves.
859. *kalen* (N) 'egg', pl. *kaleni*.
860. *kalenṭa* (N pl) 'eggs of lice'.
?861. *kalmā* (N) 'noise'.
862. *kalmā biki* (V) 'be silent'.
L863. *kāmā* (V) 'to prepare' (also used as auxiliary).
L864. *kamāy* (N) 'earning'.
L865. *kamáy* (V) 'to cause to' (auxiliary for forming causals).
L866. *kampond* (N) 'enclosed courtyard'.
L867. *kamrā* (N) 'coarse woolen blanket', cf. M. *kamble*.
L868. *kānde* (N) 'onion(s)'.
869. *kangen* (N) 'one measure of grain (roughly equal to a kilo)'.
870. *kāntā* (N pl) 'mosquitoes'.
L871. *kapācho* (N) 'cotton', cf. M. *kapus*, H. *kapās*.
L872. *kapātiŋ* (V) 'tremble', cf. K. 'tremble'.
873. *kāpliñ* (N) 'butterfly', cf. K. *kaphilij*.
874. *kapor* (N) 'winnowing basket'.
875. *kāppo* (VN) 'sleep'.
876. *kāppu* (VN) 'ritual impurity after birth'.
L877. *karamphul* (N) 'ear-ring', cf. M. *karaṇphul*.
L878. *karā* (N) 'anklet (silver)', cf. M. *kada*.
879. *karbu* (N) 'foam (on waves of sea)'.
L880. *karchi* (N) 'pitcher'.
881. *karsi* (N) see 880; cf. M. *kaḷsi*.
L882. *kāryā* (N) 'chips of wood', cf. Md. *kaḍya*.
L883. *karelā* (N) 'bitter-gourd', cf. H. *karelā*.
?884. *karjo* (N) 'cashew tree', cf. M. *kaju*; cf. *kaju-malayu*.
885. *karpi* (N) 'seller'.
886. *karpí* (V) 'to sell'.
887. *kārwēñ* (N) 'comb', cf. *kakhén*.
888. *karyom* (N) 'elder brother's wife'.
?889. *kaṭām* (N) 'silence'.
890. *kaṭhāndāri* (N) 'horsegram (split)'.
891. *kaṭhla* (N) 'armpit', Mundari *hatalaʔ*, K. *kath*^{a}*làʔ*.
L892. *kātiri* (NV) 'scissors', cf. M. *katri*.
893. *kātok-kātok* (V) 'cry of a crow'.
894. *kaṭhām* (N) 'tortoise'.
895. *kathom* (N) dialect variant of 894.
896. *kathān* (N) 'nut', cf. M. *kaṭhiṇ* 'hard (as in hard shelled nut)'.
897. *katlā* (N) 'armpit', see 891.
L898. *katthā* (N) 'catechu', cf. H. *kattha*.
899. *kāṭu* (N) 'kidney of a lamb'.
900. *kātú* (VN) 'embrace', SM. *kū̃du*.
901. *kāw* (N) 'flesh'.
L902. *kawlu* (N) 'tile', cf. M. *kaul*.
L903. *kā̃ur* (N) 'moss'.
L904. *kāwrā* (N) 'crow', cf. M. *kawla*.
905. *kāwur* (N) 'green moss', see 903.
L906. *kāyni* (V) 'to tell', cf. H. *kahānī*, Md. *kayni*.
L907. *kāykāy* 'some', cf. Md. *kaikai*.
908. *keḍā* (V) 'to feel'.
L909. *keḍe* (PP) 'towards', cf. M. *kade*.
L910. *kē̃ĩŋki* (N) 'scissors', cf. H. *kē̃ĩci*, *kā̃ĩcī*.
911. *kekhmere* (N) 'lizard', cf. K. *kekhemered*.
L912. *kelei* (N) 'tinning brass utensils', cf. M. *kalha*.
913. *kelli* (N) 'she-calf', K. *kella*, *kelli*.
L914. *kemti* 'less', cf. Hd. *kamtī*.
L915. *kendil* (N) 'hurricane lantern', cf. M. *kandil*.
916. *kepā* (N) 'louse', pl. *kepṭa*.
917. *kerchi* (V) 'to scratch'.
L918. *kere* (N) 'plantain'.
L919. *kerei* (N) 'frying pan', cf. M. *kaḍhai*.
920. *keṭṭo* (V) 'to extinguish fire'.
L921. *kināko* (N) 'wheat flour', cf. M. *kanik*.
922. *kirsā* (N) 'rich man'.
923. *kiṭā* (NV) 'winnow'.
?924. *kiṭhi* (N) 'cartridge'.
925. *kiyām* 'tomorrow'.
926. *kó* (V) 'to bring', N.B. *koken* and *košen* as future forms.

L927. *kobdur* (N) 'pigeon', cf. M. *kabutar*, K. *kobḍur*.
?928. *kocā́* (V) 'to bend'.
?929. *koccā* 'at an angle'.
930. *koḍḍu* (N) 'thumb'.
931. *kógo* (N) 'snake'.
a. *kogonjeher* (N) 'venom'.
?932. *kokhor* (N) 'hen', cf. *-koi* SM.
933. *kokõy* (N) 'ant'.
934. *kol* (N) 'wife, woman', cf. K. *koro*.
935. *kolṭa* (N) 'Nihal'.
936. *kolṭu* see 935.
937. *kol bhúto* (N) 'female ghost'.
L938. *kombā* (N) 'cock', cf. K. *kõba*, M. *komḍa*.
a. *kombā pā́to* (N) 'sacrifice of a cock'.
L939. *kollyā*^{1} (N) 'coal', cf. H. *koyla*.
L940. *kollyā*^{2} (N) 'jackal', cf. Md. *kolla*.
941. *konbiṭamāndi* (N) 'non-tribal languages', cf. M. *kunbi* 'farmer'.
942. *kondor* (V) 'to tumble'.
943. *koplyā* (N) 'hoe'.
?944. *kór* (V) 'to take away'.
L945. *kor* (V) 'to do', cf. M. *kar*.
946, *koro* (V) 'to die' (euphemistic), see 944.
947. *koṭor* (N) 'intestines'.
948. *koṭhor* (N) 'a measure of land'.
?949. *koṭṭo* (V) 'to beat, strike, pound'.
L950. *kotrā* (N) 'nest', cf. M. *kotar*.
951. *koṭṭu* see 956.
952. *kóy* see *kó*, subjunctive?
L953. *koyal* (N) 'cuckoo', cf. H. *koyal*.
954. *koyni* (V) 'to cut down a tree'.
955. *kubā* (V) 'to be intoxicated'.
956. *kubḍyā* (N) 'hunchback', H. *kubya*.
957. *kucā* (N) 'rump'.
L958. *kuḍu* (N) 'bamboo door', cf. M. *kud*.
959. *kugusu* (N) 'hair'.
960. *kuguso* see 959.
L961. *kũi* (N) 'well', cf. H. *kuā̃*.
a. *kũi dáy* (NC) 'road to the well'.
L962. *kúlā* (N) 'tiger', cf. K. *kula*.
L963. *kunbi* (N) 'farmer' (caste name among Hindu farmers).
L964. *kupurā* (N) 'clothes', cf. M. *kapḍa*.
965. *kurá* (N) 'unripe fruit'.
L966. *kurci* (N) 'chair', cf. M. *khurch*, H. *kursī*.
967. *kuri* (V) 'to apply sacred powders to the forehead'.
968. *kuruko* (N) 'ear ornament for men'.
969. *kurup* (N) 'stone'.
970. *kuṭií* (NV) 'signboard; to show the way'.
?971. *kuṭhoroy* (V) 'to cringe in fear'.

kh

972. *khá* (V) 'to remain; stay in one place'.
L973. *khājju* (V) 'to scratch'.
L974. *khāli* (Adj) 'empty', cf. H. *khali*.
L975. *khāndá* (VN) 'shoulder', cf. M. *khanda*.
L976. *khāndi* (V) 'to cut'.
L977. *khāndo* (N) 'choli-piece', cf. M. *khaṇ*.
978. *khārā* (N) 'earth', cf. 972.
a. *khārādevtā* (NC) 'earth-goddess'.
979. *khār-khori* (V) 'to clear one's throat'.
980. *khārā sārowi* (NC) 'threshing floor', cf. M. *saravṇe* 'to polish with cowdung'.
981. *khārú* (NV) 'family; herd'.
L982. *khātā* (Adj) 'sour', cf. H. *khaṭṭā*.
983. *khātāy* (N) 'raw mango', see 982.
984. *khaṭṭā* (N pl) 'words'.
?985. *khāṭi* (N) 'blacksmith'.
?986. *kheḍā* (V) 'to drive cattle; cart'.
L987. *kherido ér* (V) 'to buy', cf. *kharedi* (HM).
L988. *kherki* (N) 'window', cf. M. *khidki*.
989. *kherikama* (V) 'to pull'.
990. *kheri* (Adj) 'tight'.
L991. *kheriyān* (N) 'threshing floor', cf. M. *khaḷa*.
L992. *khéto* (N) 'field', cf. H. *khet*.
993. *khijá* (V) 'to mash together'.
994. *khilā* (N) 'parched rice'.
995. *khíjo* (V) 'to be angry', cf. Md. *khijne*.
996. *khisā* (N) 'measure of land greater than *koṭhor*'.
997. *khob insān mānso* (N) 'a proud man'.
L998. *khobo* (N) 'a lot; many'.
a. *khob diyā* (V) 'to be late'.
L/C999. *khob-jaráy* also *khojjaráy* (N) 'malaria'.
1000. *khob té* (V) 'gluttony'.
L1001. *khobor* (N) 'message', cf. H. *khabar*.
1002. *khobor mānso* (NC) 'messenger'.
L1003. *khoḍḍe* (N) 'a pit', cf. M. *khadda*.
1004. *khoḍḍo-khoḍḍo* (NC) 'place behind the knees'; see also 1259.
L1005. *khokā* (N) 'packing case', cf. M. *khoka*.
L?1006. *khol* 'valley', cf. M. *khol* 'deep'.

1007. *kholi* (N) 'black magician'.
L1008. *khorpi* (N) 'a grubbing hoe', cf. M. *khurpe*.
L1009. *khorju* (N) 'scabies', cf. M. *kharuj*.
1010. *khuḍe* (N) 'gourd'.
1011. *khulāw* (V) 'to be nice to'.
L1012. *khullā* 'open', cf. H. *khula*.
1013. *khundulur* (N) 'hump on a bull's back'.
1014. *khunu* (N) 'elbow; joint'.
1015. *khunú* (N) 'chips of wood'.
?1016. *khuri* (N) 'sole of feet', cf. M. *khur* 'sole of an animal'.
1017. *khuri-koḍḍu-ākānḍa* (NC) 'big toe'.
1018. *khuri-ko-minḍi-ǰiki* (NC) 'ankle'.
1019. *khuri tepré* (NC) 'dorsum of feet'.
1020. *khusikā jerei* (V) 'offer in gratitude'.
L1021. *khuṭā* (N) 'stump', cf. Md. *khuta*.
1022. *khutrā* (N) 'tree stump', see 1021.
1023. *khuṭo* (N), see 1021.
1024. *khuyũŋ* (V) 'to roast mahua leaves (in preparation for liquor'.
L1025. *khyālā* (V) 'to play', cf. Md. *khyaḷ*.

l

1026. *-la* suffix to form verbal participles.
1027. *lá* (pronoun) 'you' (plural).
1028. *-la* nominal suffix.
L1029. *lābhor* (N) 'rubber', cf. M. *rabar*.
L1030. *lagāṭi* 'beginning', cf. H. *lagba*.
1031. *lage jā* 'be attentive'.
L1032. *-lāgi* (PP) 'for the sake of'.
L1033. *lakari* (N) 'firewood', cf. H. *lakdi*.
1034. *laiñ* (N) 'tongue', cf. K. *lan*, NM. *alaŋ*, SM. *laʔŋ*.
a. *laiñdo kaggo* 'tongue and mouth'.
1035. *lakṭi* (N) 'a kick'.
L1036. *lambā* (Adj) 'long', cf. H. *lamba*.
1037. *lampṭāki* (N) 'oil-lamp'.
1038. *lāngi* 'web of a finger'.
L1039. *langrā* (N) 'cripple', cf. *langda* (M.).
1040. *laŋkā* (N) 'far-away place; world after death'.
1041. *lāpṭo* (V) 'to trip' (infix).
L1042. *lappo* (V) 'to bend', cf. Md. *lapne*.
1043. *laṭṭa* (PP) 'with'.
1044. *láw* (V) 'to burn'.
1045. *lāwo* (V) 'to cremate'.
L1046. *lawā* (V) 'to be tired', cf. K. *lava*?
1047. *lāyriki* 'stylishly'.
1048. *lejó* (V) 'to draw water', cf. K. *lu-*.
1049. *leñjó* (V), see 1048.
L1050. *likhātiŋ* (V) 'to write', cf. K. *likhaʔtiŋ*.
L?1051. *limbo* (N) 'neem tree'.
1052. *liwij* (V) 'to beckon'.
L1053. *lokhando* (N) 'iron', cf. M. *lokhanḍa*.
1054. *lokhod* (V) 'to stir with a spoon'.
L1055. *longo* (N) 'clove', cf. M. *lavang*.
1056. *lopot* (V) 'to be soft' (infix).
1057. *lóti* (N) 'baking plate'.
1058. *lubā* (N) 'incense'.
L1059. *lusun* (N) 'garlic', cf. M. *lasuṇ*.

m

?1060. *m* (PP) recorded only once.
1061. *mā* (V) 'to give' (rarely used).
?1062. *maccho* (Adj) 'finely woven'.
1063. *māso orki* (VC) 'to wear a mask'.
?1064. *maccho mānso* (NC) 'a healthy man'.
L1065. *mahāḍā* (N) 'a member of the mahar caste'.
L1066. *maidān jāgā* (NC) 'flat ground'.
L?1067. *mãiko(t)* (N) 'fly, mosquito', cf. *makkhi*.
1068. *makā(n)* (connective) 'but; because'.
1069. *malkā* (N) 'pea-pod'.
1070. *malkānbijo* (NC) 'bean'.
L1071. *mālik* (N) 'owner; boss', cf. H. *mālik*.
L1072. *mālum* 'known'.
L1073. *māmā* (N) 'mother's brother', cf. MH. *māmā*.
1074. *mānḍi* (NV) 'language; to translate'.
?1075. *-mandi* (PP) 'amongst'.
1076. *mānḍo* (N) 'frog', cf. Sk. *maṇḍuka*.
1077. *mándo/u* (NV) 'rain'.
1078. *mandom* (N) 'liver of a lamb'.
1079. *manê* 'we' (pl.), cf. 1580 (dual).
L1080. *mangal* (N) 'Tuesday', cf. M. *mangaḷwar*.
L1081. *mangar* (N) 'crocodile', cf. M. *magar*.
1082. *māngrya* (N) 'sweeper caste'.
1083. *mānsoki bethé* (NC) 'eunuch'.
1084. *mānṭa* (N pl) 'neighbors'.
1085. *māre* (N) 'colt'.
L1086. *-māṅ* (PP) 'because of', cf. H. *-ke mare*.
L1087. *marimāy* (NC) 'a local goddess of Hindus', cf. HM. *marimay*.
L1088. *masālā* (N) 'spices', cf. M. *masala*.

L1089. *masur* (N) 'lentils', cf. M. *masur*.
L1090. *mātā* (N) 'smallpox; also a goddess', cf. HM. *mata* 'smallpox'.
L1091. *mātāy* (Adj) 'drunk', cf. M. *matta*.
1092. *máʔaṭo* (N) 'leg; thigh', (pl.) *máʔata*.
L1093. *matta* (N) 'vote', cf. MH. *mat*.
1094. *māw* (N) 'horse'.
L1095. *māwsi* (N) 'mother's sister'.
L1096. *māy* (N) 'mother'.
1097. *māy bí* (V) 'smallpox'.
1098. *māyko* (N) 'mahua tree'.
L1099. *medā* (N) 'sheep', cf. M. *mendha*.
1100. *myego* (N) 'a god; son of Ravana', cf. Sk. *meghanad*.
1102. *mekhā* (N) 'sting of a bee', cf. M. *mekh* 'long pointed tent-peg'.
1103. *meley* (N) 'an ear of corn'.
L1104. *memsāyab* (N) 'memsahib', H. (from Urdu).
?1105. *mendān* (V) 'to clean'.
1106. *méndi* (N) 'night'.
1107. *méndi čikin* (NC) 'firefly'.
1108. *myendi jiki* (N) 'ankle bone'.
1109. *ménge* (N) 'teeth' (sg. and pl. the same).
1110. *merā* 'near'.
1111. *mér* (NV) 'play'.
L1112. *methi* (N) 'a green vegetable', cf. M. *meṭhi*.
1113. *meur* (N) 'anthill', cf. K. *bunum*.
?1114. *mhāli* (N) 'barber'.
1115. *miān* 'how much', see 1120.
1116. *micini* (N) 'a kind of fish', cf. K. *micíni*.
1117. *mijārnā* (PP) 'through'.
1118. *minḍijiki* (N) 'joint', as in *bakkoko minḍijiki* 'wrist' (*bakko* 'hand').
L1119. *milaṭiŋ* (V) 'to meet; mix', cf. K. *milatiŋ*, H. *mil-*.
1120. *mingā* 'where'; *mV-* interrogative in GRG.
1121. *minjārā* 'midst', cf. OM. *majhari*.
1122. *miŋkon* 'from where'.
L1123. *mircān* 'chilies', cf. M. *mirci*.
1124. *mirchā* (N), see 1123.
1125. *miryā* (N) 'gum (to eat)'.
L1126. *misal* (V) 'to mix', cf. M. *misaḷṇe*.
L1127. *misan* (N) 'machine', cf. Md. *misan*.
L1128. *mitiŋ* (N) 'meeting', cf. Md. *mitiŋ*.
L1129. *miṭṭhu* (N) 'pet parrot', cf. M. *miṭṭhu*.
1130. *mochor* (N) 'pestle', cf. M. *musaḷ*.
1131. *mokh(a)ne* (N) 'knee' ('elbow' for some).
1132. *molki jām* (V) 'to change money'.
1133. *momāsānā* (N) 'master'.
L1134. *momo* (N) 'krait' (a snake), cf. K. *momo*.
1135. *monḍā* (N) 'heel'.
L1136. *mongo* (N) 'green gram', cf. Md. *mũng*.
L1137. *mongorā* (N) 'jasmine', cf. M. *mogra*.
1138. *mongyā* (N) 'tattooer'.
1139. *monjo* (Adj) 'silent'.
1140. *monjjer* (V) 'to bless', cf. M. *munja*.
1141. *mor* (N) 'thorn'.
1142. *morko* (V) 'to pass wind'.
1143. *morku* (N) 'vagina'.
1144. *moṭh* (num) 'three'.
a. *moṭhā din* 'day after tomorrow'.
L1145. *motoru motur* (N) 'car', cf. M. *motar*.
1146. *mowri*.
L1147. *muār* (N) 'mouth', cf. K. *muar*.
1148. *mũikku* (N) 'wedding necklace of black beads'.
1149. *mujri* (V) 'to work as a laborer'.
1150. *mukhā* (N) 'a measure of two arm lengths'.
1151. *mukhyā* (N) 'the man who foretells when possessed'.
1152. *mulākā* (N) 'region', cf. M. *mulukh*.
1153. *mundāl pālšo* (N) 'child of former husband'.
L1154. *mundi* (N) 'ring', cf. Md. *mudi*.
L1155. *mung* (N) 'green gram', cf. M. *mũng*.
?1156. *mungukā* (N) 'dirty', cf. K. *kumuka*.
1157. *munguso* (N) 'mongoose'.
1158. *murkitij* (N) 'small insects', cf. OM. *murkute*, (pl.) *murkiṭṭa*.
L1159. *musā* (N) 'mustache', cf. *mučh*, H. *mũc*.
L1160. *muskil* 'difficult', cf. H. *muśkil*.
L1161. *muṭhi* (N) 'fist', cf. M. *muṭh*.

n

1162. *-nā* (PP) nominal suffix for 'by' and 'possession'.
1163. *nānā* 'what'.
1164. *nāni* 'who'.
L1165. *nakkho* (N) 'nail', cf. M. *nakh*.
L1166. *nām kāmā* (V) 'to earn a name (fame)'; borrowed idiom from H. *nām kamana*.
1167. *nānkā* 'of what sort?', see 1163.
L1168. *nāli* (N) 'small stream', cf. H. *nali*.
?1169. *nālku* (num) 'four' (Dravidian source), also *nhālku*.
L1170. *nāngar* (N) 'plough', cf. M. *nangar*.

1171. *nāpyam* (N) 'mother's brother's wife'.
1172. *napyom* (N) 'mother-in-law', see 1171.
1173. *napyāv* (N), see 1172.
1174. *nāniki* 'how', see 1163, 1164.
L1175. *nāngo* (N) 'cobra', cf. Md. *nang*.
?1176. *nārā* (N) 'throat', cf. M. *narde* 'Adam's apple'.
1177. *nārāki bá* (V) 'choke on something'.
1178. *nārāpongā* (NC) 'trachea'.
L1179. *nāri* (N) 'umbilical cord', cf. M. *naḷ*.
L1180. *nārini* (N) 'breakfast', cf. Md. *nyahri*.
L1181. *noromā* (Adj) 'soft', cf. M. *naran*.
L1182. *nāsā* (N) 'veins', cf. M. *nasa*.
1183. *nāthā* (N) 'nose-ring', cf. M. *nath*.
L1184. *naṭṭā* (N) 'snub-nosed', cf. M. *nata*.
L1185. *navā* (Adj) 'new', cf. M. *nava*.
1186. *nāwki* 'in what way', see 1163.
1187. *navār* (N) 'python'.
1188. *nawā* 'of what type', cf. *hūwa* 'of this type'.
1189. *nāvay* 'why'.
L1190. *nāy* (N) 'dog' (Dravidian).
L1191. *ne* (PP) 'instrumental', cf. M. *ne*.
1192. *nyé* (pronoun) 2nd person sg.
L1193. *nesnā* (V) 'to dress', cf. M. *nesṇe*.
1194. *néto* (N) 'ash'.
L1195. *newdi* (V) 'to offer to gods', cf. M. *naivedya*.
1196. *néygo* (N) 'scorpion'.
a. *néygon cutiñ* (N) 'sting of a scorpion'.
L1197. *nhanāṭa* (N pl) 'children'.
1198. *ni* (PP) 'for'.
1199. *nidir* (N) 'white ant'.
L1200. *nimbo* (N) 'neem tree'.
L1201. *nimbu* (N) 'lime'.
L1202. *nindo* (V) 'to weed', cf. M. *nindṇi*.
L1203. *nindo* (N) 'sleep', cf. Md. *nind*.
1204. *niṭṭi* as in *āwār-nitti* see under *āwār*.
1205. *niṭṭo* (V) 'enter'.
1206. *niṭṭhó* (V) 'to sharpen'.
L1207. *niwrāwēn* (V) 'to sift grain', cf. M. *nivaḍ*.
1208. *nyeko* (pronoun) 2nd person plural.

o

All words are optionally available as beginning with *wo*/*vo*. There is no contrast */wo/ ~ /o/.

1209. *-o* added to roots to form base when confirming to syllable formation rules.
L1210. *ocol* (V) 'to lift', cf. M. *ucal*.
1211. *oḍā* (N) 'temple', in N.M. 'house'.
1212. *odow* (N) 'she-buffalo'.
?1213. *ochoni* (V) 'to boil (as water)', e.g., *joppo ochonika* 'water boils'.
?1214. *óhan* (N) 'mortar (with pestle'.
1215. *ojāba* (V) 'plant'.
L1216. *okhar* (N) 'plough', cf. Md. *okhar*.
1217. *okki* (V) 'to keep; store'.
1218. *okúd* (V) 'to hiccough'.
1219. *okum* 'then'.
L1220. *olā* (V) 'to be wet', cf. M. *ola*.
?1221. *olān* (N) 'cooked dhal', cf. Md. *aḷaṇ*.
1222. *olki* (V) 'to have fungus'.
1223. *opónā* (N) 'day'.
1224. *órā* (N) 'air, wind'.
a. *orāŋdo* (N) 'wind'.
1225. *oriḍ* (V) 'to strip'.
1226. *oró* (N) 'millet, jawar'.
L1227. *orṭā* (V) 'to turn', cf. M. *ulta*.
1228. *orṭāk* (V) 'to be lost'.
L1229. *osāri* (N) 'deserted place', cf. M. *osadi*.
1230. *oṭhā* (N) 'lip'.
1231. *oṭho* (N) 'chin'.
1232. *othārā* (N) 'dwarf'.
1233. *oṭhyā* (N) 'harrow'.
1234. *oṭi* 'that'.
1235. *oṭṭi* (V) 'to pull out'.
1236. *owāri* (N) 'son's wife; younger brother's wife', cf. H. *bouhaṅ*.
L1237. *oyjā* (VN) 'load; to carry a load', cf. K. *oyja*.
1238. *oyni*[1] (V) 'to weed'.
1239. *oyni*[2] (V) 'to beg'.

(Not alphabetically ordered) additional *-o-* words (in another dialect).

L1240. *worāri* (N) 'wedding guest', cf. M. *worari*.
1241. *worāri pakkā* (N) 'engagement'.
1242. *wortāy* (V) 'to arrive'.
1243. *wonci* (V) 'to beg'.
1244. *wokki* (V) 'to arrange'.
1245. *woḍi* (N) 'carpenter'.
L1246. *wos* (N) 'dew', cf. H. *os*.
1247. *woṭ* (N) 'vote', cf. H. *vot*.
1248. *wondor* (N) 'biceps'.

p

1249. *páḍā* (V) 'to kill'.
1250. *págo* (N) 'tail'.
L1251. *pagrí* (N) 'head turban', cf. M. *pugdi.*
L1252. *pahāḍ* (N) 'mountain', cf. H. *pahad.*
L1253. *pāhuncāri* (NV) 'welcome', cf. M. *pahuncar.*
L1254. *pāhun* (V) 'to be a guest', cf. M. *pahuṇa.*
1255. *pāylu* (N) 'young parrot'.
L1256. *pakkā* (Adj) 'definite', cf. H. *pakka.*
1257. *pākin* (N) 'hen' (OSk).
1258. *pakhorā* (N) 'upper arm'.
?1259. *pakhorāko khoḍḍo-khoḍḍo* (NC) 'armpit'.
?1260. *pakhoroko danḍi* (N) 'forearm'.
?1261. *pakhorā aṭṭu* (V) 'to stretch oneself'.
L1262. *pākho* (N) 'wing'.
L1263. *palā*^{1} (N) 'leaves of a tree', cf. M. *pala.*
L1264. *palā*^{2} (N) 'shop in the weekly market'.
L1265. *pān* (N) 'betel leaf'.
L1266. *pānco* (N) 'village elder(s)', cf. M. *panca.*
L1267. *pānḍo* (V) 'to cross over', cf. M. *phandṇe.*
L1268. *pãnc* (num) 'five'.
L1269. *pānder cunā* (N) 'gravel'.
L1270. *pāndher* (Adj) 'white'.
L1271. *pankhā* (N) 'fan'.
L1272. *pāpi* (Adj) 'sinful, guilty', cf. M. *papi.*
L1273. *pár* 'everything', cf. M. *par.*
1274. *parā̃y* (N) 'river', cf. *konkani parhya.*
?1275. *pasalā* (Adj) 'second; other'.
1276. *pachālk* (V) 'to itch'.
?1277. *pasālā* (N) 'joint family paraphernalia'.
1278. *pát ará* (V) 'to wait for someone'.
1279. *pātā́* (V) 'to attack'.
L1280. *patār*^{1} (Adj - V) 'thin', cf. M. *pataḷ.*
1281. *patār*^{2} (V) 'to take clothes down from the line'.
1282. *pātāri* (Adj) 'dry'.
1283. *patār-kamā̃y* (V) 'to dry in the sun'.
1284. *paṭāri boy* (NC) 'straw, dry grass'.
L1285. *paṭel* (N) 'chief'.
L1286. *paṭeri* (Adj) 'thin; dilute', cf. 1280.
1287. *paṭhā* (Adj) 'elder; old'.
1288. *paṭhyā* (N) 'aged man'.
1289. *-pāti* (PP) 'inside', e.g.,

cán	*gen̈pati*	*bākijerei*
'fish'	'fishhook-in'	'catch-ready'

1290. *páto* (V) 'to come, to approach'.
1291. *páto sāl* 'next year'.
1292. *paṭṭo* (V) 'come', see 1290.
L1293. *pāwi* (N) 'flute', cf. M. *pava, pawri.*
L1294. *pāwlyā* (N) 'flute player'.
L1295. *pāwso* (N) 'shower of rain', cf. M. *paus.*
L1296. *pāyri* (N) 'step', cf. M. *payri.*
1297. *peḍe* (V) 'to be in trouble; suffer'.
1298. *péjo* (V) 'to step on'.
1299. *pejāgo* (V) 'to elope', e.g., *itkel pejgi* 'the two eloped'.
1300. *pejāk-kāmāy* (V) 'to abduct'.
1301. *pejikoenkāmā* (V) 'to drive away (cattle)'.
1302. *péku-jeré* (V) 'to hide in ambush'.
1303. *péko* (N) 'cave'.
1304. *pephryā* (N) 'saliva', cf. M. *phepre.*
L1305. *pénḍ(i)ri* (N) 'calf of the leg', cf. M. *pindri.*
1306. *pyeŋ* (N) 'head'.
L1307. *pensil* (N) 'pencil'.
1308. *péráy* (V) 'to stay'.
?1309. *peri* (N) 'phalanx'.
1310. *perto* (V) 'to squeeze'.
1311. *péte* (V) 'to sit'.
1312. *peté* (V) 'to alight from sky'.
1313. *petel* see *patel.*
?1314. *petek* (V) 'to break a thread'.
1315. *petekkāmā* (V) 'to tear clothes'.
1316. *péto* see *péte.*
1317. *petto* (V) 'wring'.
1318. *petekāmā* (V) 'to make to sit'.
1319. *pewrikere* (N) 'plantain'.
?1320. *pewṅki ajni* (V) 'to vomit (bile)'.
?1321. *pyetako* (Adj) 'torn'.
1322. *peyko peyko* (N) 'stalking gait'.
1323. *piji bérto* (N) 'elder brother-in-law'.
L1324. *pillāto* (N) 'plate'.
L1325. *pin* (N) 'hairpin', cf. M. *pin* (through English).
1326. *pindá* (V) 'to jump'.
1327. *pindiri gurohā* (N) 'shin'.
1328. *pírjo* (N) 'girl; daughter'.
1329. *pitrāki goyi* (N) 'lead bullet'.
1330. *piu* (N) 'a bird locally called *pawsa* and said to be giving notice of rain'.
1331. *piviṭo* (N) 'hill, field'.
?1332. *piyā*^{1} (N) 'spleen of a lamb'.
1333. *piyā*^{2} (V) 'to beckon'.

L1334. *piyeco* (N) ‘key of a lamp’.
L1335. *polkã* (N) ‘blouse’.
1336. *polor* (V) ‘to exist’.
1337. *pongā* see *narā pongā*.
1338. *popā́* (N) ‘hole’.
1339. *-popā* see *cón popa*.
?L1340. *poppāṭi* (Adj) ‘strong wind’, cf. *phuphata*.
1341. *poppá* (V) ‘to bore a hole’.
1342. *popo* (N) ‘nipple’ (‘belly’ in another dialect).
1343. *popo ér* (V) ‘to suffer from dysentery or cholera’.
1344. *poporkā* (Adj) ‘dirty’, *-ka* Korku suffix.
1345. *poporā* (Adj) ‘dark’.
1346. *pori mānso* (N) ‘male orphan’.
1347. *pori pírjo* (N) ‘male orphan’.
L1348. *potā* (N) ‘sack’, cf. M. *pota*.
L1349. *powrā* (N) ‘bucket to draw water’, cf. Md. *poyra*.
?1350. *powāri* (N) ‘conch’, cf. M. *powḷi* ‘coral’.
1351. *poy* (N) ‘bird’, (pl.) *poyṭa*.
1352. *poye* (N) ‘feather’, (pl.) *poyeṭa*.
L1353. *puch* (V) ‘to wipe away’.
1354. *punam* (N) ‘full-moon day’.
L1355. *pũnjā* (V) ‘to make a heap’, cf. M. *punjka* ‘a heap, cluster’.
L1356. *punye* (N), see 1354.
L1357. *puju* (V) ‘worship’, cf. M. *puja*.
L1358. *purā* (Adj) ‘total’.
L1359. *purá* (V) ‘to be complete’.
1360. *púra* (V) ‘to send’.
1361. *purlā sāl* (NC) ‘last year’.
1362. *púsi* (V) ‘to erase’.
1363. *pusu* (V) ‘to change money’.
1364. *pusuy* (V) ‘to escape from pursuit’.
1365. *putá* (V) ‘to be unfertile’.
1366. *pútā* (V) ‘to bloom’.
1367. *putkiri-ja* (N) ‘measles, chicken-pox’.

ph

1368. *phaḍḍu-phaḍḍu* (V) ‘to splash water’.
1369. *phāri-kāmāy* (V) ‘to divide property’.
L1370. *phārkati* (N) ‘divorce’.
L1371. *pharsi* (N) ‘tile on the floor’.
L1372. *phāṭako* (N) ‘gate’, cf. M. *phāṭak*.
L1373. *phāwra* (N) ‘spade’, cf. M. *phāvḍe*.
L1374. *phejer* (N) ‘morning’.
L1375. *phellyā* (N) ‘groundnut’.
1376. *phendrā* (N) ‘vine’.
1377. *phér* (V) ‘to take a picture; draw’.
1378. *pherā* (V) ‘to search’.
1379. *pherātẽñ* (dual noun) ‘hen and cock’.
1380. *phernikā* (dual noun) ‘cow and bull’.
L1381. *phetkā* (N) ‘firecracker’.
L1382. *phikā* (Adj) ‘mild, pale’.
1383. *phikāy* (V), see 1382, ‘to be pale’.
1384. *phiki oyni* (V) ‘to beg’, cf. M. *bhik*.
1385. *phittākamāy* (V) ‘change of seasons’.
L1386. *phopsā* (Adj) ‘soft’.
L1387. *phór* (NV) ‘fruit; to bear fruit’.
L1388. *photre* (N) ‘chilies’.
1389. *phúl* (NV) ‘flower; to blossom’.
1390. *phul* (N) ‘bridge’.
1391. *phunṭol-phunṭol* (N) ‘rattle (for babies)’.
L1392. *phuṭānā* (N) ‘roasted horsegram’.

r

1393. *rābān* (N) ‘cold’.
1394. *raccho* (N) ‘honey; any juice’.
?1395. *rāggi* (V) ‘to stand in a line’, cf. M. *raŋ* ‘queue’.
1396. *rāgo* (N) ‘Adam’s apple’.
1397. *rakhom* (V) ‘wait’.
L1398. *rām rām* (N) ‘greetings’, cf. Md. *ramram*.
L1399. *rānḍ* (N) ‘prostitute’.
L1400. *rāndā* (V) ‘to cook’, cf. M. *randhṇe*.
L1401. *rangā* (N) ‘color’.
1402. *rangó* (V) ‘to turn brown through overcooking’.
1403. *rāngo-kāmāy* (V) ‘to dye in color’.
1404. *rāngiṭṭa* (N) ‘molars’.
1405. *rāon* (N) ‘Ravan, a mythological demon of Hindus and a god of Nihals’.
1406. *rāphid* (V) ‘to blink’.
L1407. *rāsto* (V) ‘to blink’.
1408. *rāwanbí* (V) ‘to die’.
L1409. *rãwa* (N) ‘parrot’.
1410. *raymonyā* (N) ‘a wild thorny bush’.
1411. *re* personalizer.
1412. *redigo* (N) ‘radio; tape-recorder’.
1413. *rekki* (V) ‘to crush’.
1414. *reweḍi* (Adj) ‘tight’.
1415. *ritáy* (V) ‘to give birth’.
1416. *riṭhā* (N) ‘soap nut’.
1417. *robān* see *raban*.
L1418. *rogā* (N) ‘disease’.

L1419. *rogi* (N) 'diseased; sick man'.
L1420. *rojoko* 'everyday'.
L1421. *rojoka* see 1420.
1422. *rangāyo* see 1402.
1423. *roṭā* (Adj) 'ripe'.
L1424. *rupiya* (N) 'rupee', cf. M. *rupaya*.

s (*š* more often, occasionally *s*)

1425. *sa* (num) 'six'.
L1426. *sababaki* 'because of'.
1427. *sabu* (NV) 'soap'.
L1428. *saco* (N) 'breath'.
L1429. *sacco* (N) 'true'.
L1430. *sadbhai* (N) 'brother-in-law'.
L1431. *sadu* (N) 'medicine man'.
L1432. *sagai* (N) 'engagement'.
L1433. *sagāy* (N) 'family (vert. extended)'.
L1434. *sagākkā* (N) 'descendants'.
L1435. *sagarā* 'all of it'.
L1436. *sajā* (N and V) 'true'.
L1437. *sajā* (N and V) 'real' (as opposed to magical).
L1438. *sajākā* 'pleasant'.
L1439. *sokārikā* 'early'.
1440. *sākhob* (V) 'to sweep'.
L1441. *sakkhar* (N) 'sugar'.
1442. *sokorā* (N) 'bread'.
L1443. *sākryā* (N) 'sugar'.
L1444. *salāy* (N) 'sala tree'.
L1445. *salāi*, cf. 1444.
L1446. *sáli* 'chaff'.
L1447. *sáli* 'bark of a tree'.
L1448. 'husband; term of reference'.
L1449. *sāndā* (N) 'joint'.
L1450. *sandako* (N) 'box', cf. H. *sanduk*.
L1451. *sarkār* (N) 'government', cf. H. *sarkar*.
?1452. *sāvan* (N) 'plain on the hill; plateau'.
1453. *sāvan* (N) 'of even keel'.
1454. *sammāki pātokā sāl* 'year after next'.
L1455. *sapā* (Adj - V) 'clean', cf. MH. *saph*.
L1456. *sapā-kamāy* (V) 'curse', cf. M. *šap*.
1457. *sarkā ubā* (Adj) 'vertical'.
1458. *sarikā* (V) 'to tidy up'.
L1459. *sarrak* (N) 'road', cf. M. *sadak*.
L1460. *sātho* (PP) 'accompanying with', cf. H. *sath*.
L1461. *sāvto* (Adj) 'step relation', cf. Md. *savta*.
L1462. *sāybo* (N) 'sahib'.
L1463. *sāykal* (N) 'bicycle'.
1464. *sé* 'yesterday'.
L1465. *sekó* (V) 'sneeze', cf. M. *sink*.
?1466. *selki* (N) 'reflection in water'.
L1467. *semru* (N) 'nasal mucus', cf. M. *sembud*.
L1468. *senā* (N) 'cow-dung'.
L1469. *sendā* (N) 'point; top', cf. M. *šenda*.
1470. *sendā ākhanḍā* (N) 'little finger'.
1471. *sendā ānguṭha*, see 1470.
L1472. *sepu* (N) 'fennel'.
1473. *seprā* (N) 'rib'.
L1474. *seri* (N) 'goat', cf. M. *šeḷi*.
1475. *serto-kamāy* (V) 'to give a feast'.
L1476. *sewāĩ* (N) 'macaroni', cf. Md. *sevāĩ*.
1477. *siḍu* (N) 'mahva wine'.
1478. *sikā*^{1} (N) 'plastic clay'.
1479. *sikā*^{2} (N) 'pubic hair'.
L1480. *sikār* (V) 'to blow nose'.
L1481. *sikāri* (N) 'hunter'.
L1482. *sikó* (V) 'to learn', cf. M. *šikṇe*.
L1483. *simburu*, see *semru*.
L1484. *sikret* (N) 'cigarette'.
1485. *simélā* 'torn'.
L1486. *sind* 'the sindi tree (palm)'.
L1487. *singi* (N) 'horn(s)'.
L1488. *sinwar* (N) 'Saturday'.
L1489. *sipi* (N) 'tailor'.
L1490. *sipnā* (N) 'teak'.
1491. *siprā* (N) 'rib'.
L1492. *sirā* (N) 'blood vessel; veins'.
L1493. *sirāp* 'only', cf. H. *sirph*.
L1494. *sitā* (N) 'dog'.
L1495. *sitāphal* (N) 'custard apple'.
L1496. *siwnā* (N) 'dream'.
1497. *sobojeré* (V) 'to pierce'.
L1498. *sobbi* (N) 'together'.
1499. *sojākā* 'well, good'.
1500. *sokko* (V) 'to rinse clothes'.
1501. *sokorā* (N) 'bread'.
1502. *sommār* (N) 'Monday'.
L1503. *songo* (N) 'pretense'.
L1504. *sonā* (N) 'gold'.
L1505. *sonān* (N) 'goldsmith'.
1506. *sonu* (N) 'black-eyed beans'.
1507. *soso*.
L1508. *soṭā* (Adj) 'small'.
?1509. *sowādi* (V) 'to rot'.
L1510. *sūĩ* (N) 'needle'.
L1511. *suīni* (N) 'midwife', cf. M. *suiṇ*.

1512. *sukurār* (N) ‘Friday’.
1513. *sunār*, see 1505.
1514. *sunḍu* (N) ‘pod for the beans’.
L1515. *supāri* (N) ‘betel nut’.
L1516. *suri* (N) ‘knife’, cf. M. *suri*.
L1517. *suru* (V) ‘to start’, cf. M. *suru*.
L1518. *suṭṭi* (N) ‘holiday’.
L1519. *suto* (N) ‘thread’.

ch: additional list

1520. *chabbāl* (N) ‘iron hoe’.
1521. *chāgā* (N) ‘a variety of grass (thorny)’.
1522. *chāgo* (N) ‘small stones’.
1523. *chākāw* (V) ‘sweep’.
L1524. *chāli* (N) ‘bark’.
?L1525. *chāmbor* (N) ‘female animal’.
1526. *cháʔar* (V) ‘to swell’.
1527. *chārbān* (N) ‘trellis’.
1528. *chāre* (V) ‘save; cure’.
L1529. *chāti* (N) ‘breast; chest’.
L1530. *chāṭo* (N) ‘fast’.
1531. *chebela* ‘third’.
1532. *chellā* (N) ‘lizard’.
1533. *chemic* (N) ‘chicken dung’.
1534. *chepiyā* (N) ‘variety of grass’.
1535. *chenā-jhāpon* (N) ‘mushroom that grows on cow dung’.
1536. *chenki* (V) ‘to go mad’.
1537. *cherekkā* (N) ‘pus’.
1538. *chiḍu* (N) ‘wine’.
1539. *chikā* (N) ‘pubic hair’.
1540. *chillār*.
1541. *chim* (V) ‘to sew’.
1542. *chiwār* (N) ‘clay’.
1543. *chobo* (V) ‘to stab’, cf. *sobojerei*.
1544. *chondevo* (N) ‘pus’.
1545. *choŋ koro* (N) ‘masked man’.
1546. *choroḍi* (N) ‘mould on food’.
1547. *chũi*, see *sũĩ*.
1548. *chuluk* (V) ‘to kindle’.
1549. *chunco* (N) ‘a vegetable’.
1550. *chunḍu* (N) ‘bean’.
1551. *chu* (V) ‘to spoil’.
1552. *chyolāy jhāpon* (N) ‘a small mushroom growing on damp soil’.

t

1553. *ṭā*¹ (PP) ‘with regards to him’.
1554. *ṭā*² plural morpheme.
?1555. *ṭāgā* prefix to nouns, ‘because of’.
1556. *tagāli pāktó* (N) ‘collar bone’.
1557. *tágo* (Adj) ‘fierce’.
?1558. *tágo caini* ‘one after the other’.
1559. *tagā jhagrā* ‘violent quarrel’.
1560. *tahāʔāri* (N) ‘pregnancy’.
L1561. *takādo* (N) ‘strength’, cf. M. *takad*.
L1562. *ṭakkarā* (V) ‘challenge’.
?1563. *ṭākkhocu* (N) ‘coitus’.
L1564. *tālki* (N) ‘rhythm of the drum’.
1565. *talāri(re)* (num) ‘third’.
1566. *ṭalpono* (num) ‘fourth’.
L1567. *tamākoʔo* (N) ‘tomato’.
L1568. *tambyā* (N) ‘brass’.
1569. *tāmku* (N) ‘tobacco’.
1570. *tāndur* (N) ‘rice’, also ‘cooked rice’.
1571. *tangoi* ‘stretched’.
1572. *tāpne*, see *jiki tā́pne* (792).
1573. *táro* (V) ‘throw’.
1574. *tarāke* (N) ‘bits of broken glass’.
L1575. *tarāso* (N) ‘cruelty’.
1576. *tār* (V) ‘pick out’.
L1577. *tarsyā* (N) ‘hyena’, cf. M. *taras*.
?1578. *taṭhyā* (N) ‘herdsman’; also, Korkus are called *taṭhyā* by the Nihals.
L1579. *tejo* (Adj) ‘sharp’.
1580. *ṭéko* (pronoun) ‘we’ (dual), cf. 1079.
1581. *té* (V) ‘eat’ (SM.[GRG] *tej-* ‘to serve food’).
1582. *-ṭel* dual suffix for nouns.
L1583. *tel* (N) ‘oil’, cf. H. *tel*.
1584. *ṭepré* (N) see *khuri tepré*.
1585. *terei* (N) ‘waterfall’, cf. *paruy*.
1586. *ṭevre* (N) ‘lips’.
1587. *ṭẽyā* (N) ‘brother-in-law’, cf. K. *tiỹã ~ teỹã*.
L1588. *thuk* (N) ‘spit’.
1589. *-ṭi* (PP) ‘with regards to him’.
L1590. *ṭíka* (N) ‘kumkum on the forehead’.
1591. *ṭimki* (N) ‘a kind of drum’.
L1592. *-tiŋ* verbalizer, cf. K. *-teŋ ~ -tiŋ*.
1593. *tircha* (N) ‘squint-eyed’.
1594. *tirochaki arā* (V) ‘to look squintingly’.
1595. *tiriñchimá* (V) ‘to borrow’.
1596. *tisrā din* ‘day before yesterday’.

1597. *titimāri* (N) ‘silk’.
L1598. *titiw* (N) ‘lapwing stilt’, cf. M. *titvi*.
L1599. *tiwār* (NV) ‘festival; to be festive’.
1600. *ṭiwi* (V) ‘to carry in hand’.
1601. *ṭó* (N) ‘ear of corn’.
1602. *ṭokā* (V) ‘to place on’.
1603. *tól* (N) ‘skin’.
1604. *ṭongre* (N) ‘knee’.
L1605. *toplā* (N) ‘basket’, cf. M. *topla*.
1606. *topo* (N) ‘slope’.
1607. *topti* (V) ‘to glue, stick’.
1608. *torá* (V) ‘taste hot like a pepper’.
1609. *toroʔo* (N) ‘a woven carpet’.
1610. *toṭā* (N) ‘clan name’.
L1611. *ṭoblyā* (N) ‘stammerer’.
1612. *ṭoṭor* (N) ‘grasshopper’.
1613. *ṭuguṭ* (V) ‘to ripen’.
1614. *ṭugiṭi* ‘ripe’.
?1615. *tukrān* ‘partial’.
L1616. *turi* ‘gram’.
1617. *turi dān* ‘split yellow gram’.
1618. *ṭutini* (V) ‘to carry water’.

th

L1619. *ṭhagātiŋ* (V) ‘to cheat’.
1620. *thanḍā-thunḍā* (N) ‘leprosy’.
L1621. *thandāyki* ‘drink with bhang’.
L1622. *thāpṭi* (N) ‘clap’.
L1623. *thārā* (N) ‘plate’.
1624. *thári* ‘again’.
1625. *ṭhekri* (N) ‘forehead’.
1626. *thyén* (V) ‘to stay’.
1627. *ṭhen* ‘much’.
1628. *ṭhen māndí* (V) ‘to chatter’.
1629. *thénded* (N) ‘moon’.
1630. *thénded giryān* (N) ‘lunar eclipse’.
L1631. *thendi* (N) ‘cold weather’.
1632. *ṭhikin*, see 1625.
1633. *ṭhikri golāy* (V) ‘wrinkle’.
1634. *ṭhikin tól golāy* (V) ‘to knit one’s brows’.
L1635. *thorā* (V) ‘to be less’.

u

1636. *ubā* (N) ‘straight’.
1637. *ubā dāndi pakto* (N) ‘vertebral column’.
1638. *ucā* (N) ‘cliff; tall’.
1639. *ucā oṭhrā* ‘rough plain’.
1640. *uḍi* (V) ‘to rub’.
1641. *uḍiḍone* ‘dragging’.
L1642. *ujo* (N) ‘light (not darkness)’.
1643. *ugār* (V) ‘to open’.
1644. *ugār jhāpnā* (N) ‘uncover’.
1645. *ukrā* (VN) ‘exaggerate’.
1646. *ukhuṅ* (V).
1647. *ulṭāy* (V) ‘to break’.
L1648. *ulṭā* (VN) ‘opposite; reverse’.
1649. *ulta pāv* ‘fallen’ (idomatic).
L1650. *umber* (N) ‘age’.
1651. *ulusu* (V) ‘call of a cat’.
L1652. *umān* (V) ‘to measure; to count’.
1653. *ungāy* (V) ‘live’.
1654. *urā* (V) ‘to rise; get up’.
L1655. *urātiŋ* (V) ‘spread’.
1656. *ú* (V) ‘kindle’.
1657. *urí* (V) ‘to use; apply’.
1658. *uruḍ* (V) ‘to brush (teeth)’.
L1659. *urphā* ‘alias’.
1660. *uru*, see *ú* — *ú* and *uru* are dialect variants; this is historically significant, *-ru* may be isolable.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali Lexicon: Supplement I
Words Collected by Bhattacharya and Konow

Compiled by Hal Fleming

This supplemental lexicon contains material collected by the earlier field-workers, Sten Konow (Grierson 1906) and Sudhibhushan Bhattacharya (1957). This earlier material is reprinted here only insofar as (a) Mundlay's lexicon does not include the word, or (b) there is a significant phonetic or semantic difference between Mundlay's form (M) and/or Konow's (K) and/or Bhattacharya's (B); e.g., *pyeŋ* (M), vs. *peṅg* (K), vs. *pẽ̄y* (B) 'head'. Numeral words are listed separately (see Supplement II).

Items are listed in Roman alphabetical order, except that aspirates follow non-aspirates (e.g., *bh* after *b*), and retroflex follows dental. Notes to some items have been appended by Hal Fleming. Some kinship terms are abbreviated as follows: Br = brother, El = elder, Fa = father, Hu = husband, Mo = mother, Si = sister, So = son, Wi = wife, Yo = younger.

1. *accha* (B) 'good'. Bhat[tacharya] says cf. Hindi.
2. *achuḍ-* (B) 'to hand something' (M *asudi*).
3. *aḍek-* (B) 'to burn (intr.), burn by itself' (M *aḍé*).
4. *akhanḍi* (B) 'finger'. Bhat says cf. Skt. *aṅguṣṭha* 'thumb'. That is an accidental similarity, methinks (M *akhaṇḍā*, *akhuṇḍa*).
5. *an* (B) 'other'. Bhat says cf. Skt. *anya-*. Methinks not necessarily, what with [n] in demonstratives and such being so common (M *ane*).
6. *anci-* (B) 'to select' (M. *anchi*).
7. *aŋgarako* (B) 'shirt' (M *angārko* 'upper garment of a man').
8. *aŋgluij-* (B) 'to bathe'. Awfully long root for a simple verb (M *angulij*).
9. *ānti* (B) 'for'.
10. *āpo* (B)(K) 'fire' (M *ápó* 'wood, to be lit').
11. *ara* (B) 'to see' (M 'to examine closely').
12. *arḍu/aḍḍo* (B) 'tree'. Bhat says cf. Parengi *ara*, Skt. *dā́ru*. Sono dubbio because the Sanskrit form is also cognate with *tree* in English and Swedish *trä(d)*. (M *aḍḍo*.)
13. *awalka* (K) 'good, better, best'. All the same. (M *awol* 'good'.)
14. *bā/eba/aba* (B) 'father' (M *abā*).
15. *baba* (B) FaElBr, FaSiHu. Bhat says cf. Bengali *babe* 'father'. Yeh, and half of the rest of the world. Sono dubbio. (M MoFa.)
16. *bachye* (B) 'younger' (cf. M *base* 'small').
17. *baḍḍi* (K) 'bull'.
18. *bai* (B) ElSi, (K) Si (M 'woman, ElSi').
19. *baṭe/bāṭe* (B) 'now'.
20. *bekki* (B) 'to reap' (M 'to sow, to reap', also *bétki*).
21. *berko* (B); *berkū* (K) 'cat' (M 'frog' in one dialect, 'tomcat' in another).
22. *bi* (B) 'also'. Adverb. Bhat says cf. Korku; Hindi *bhī*.
23. *bica* (K) 'why?' (M *bicá* 'to ask a question').
24. *birṭom* (B) HuElBr, WiElSi, 'father-in-law' (M FaElBr, WiElSi, 'father-in-law').
25. *boko/bokko* (B) 'arm'. It is 'hand' to Grierson(-Konow). (M *bakko* 'palm of the hand'.)
26. *bologo* (B) 'bear' (animal). Bhat says cf. Skt. *bhalluka*. Amen to it. Sanskrit was not necessarily the source of loan words into Nihali in ancient times; a resemblance to it may just mean a source in Old Indic and if, like this word here, the Indo-European derivation is not so clear, then the loan may have gone from Nihali into Old Indic. Good to bear in mind. (See Mundlay's lexicon, no. 237.)
27. *bommoki* (B) Br (M *bo(m)moki* 'siblings').
28. *boṇḍe* (K) 'near'.
29. *bōy* (B) 'grass, fodder; name of a Nihali clan' (M *bóy* 'grass, reed, weed, etc.').
30. *bhaga-dhawa* (K) 'fear' (M *bhāgā* 'big, much', *dhāvā* 'distant').
31. *bhaga may* (B) FaElBrWi, MoElSi, or aunt with seniority. First form = 'big' (M 'grandmother').

32. *bhagiya* (K) ‘slave’.
33. *bhanja* (B) SiSo. Bhat says it is from Hindi.
34. *bhāwḍi* (K); *bhavri* (B) ‘back (of body)’ (M *bhāwri*).
35. *bhītarke* (K) ‘down’.
36. *bhūt* (K) ‘devil’, ergo ‘Satan’. Cf. Juang *bhuta*, K says. (M *bhuto* ‘male ghost’.)
37. *caini* (K) ‘before’ (location) (M *cayni* ‘first’ [num]).
38. *cakhaw-* (B) ‘to sweep’ (M also *chākāw*).
39. *carkaḍ* (B) ‘waist’ (M *carkhaḍ*).
40. *cekoṭo* (B) ‘axe’ (M *cekoṭo* ‘to chop meat’, *cekṭo* ‘knife’).
41. *cyō-* (B) ‘to urinate’ (M *có*).
42. *chaka-* (B) ‘to ascend, climb up’ (M *caká*, *cāko*).
43. *che* (B) ‘yesterday’. Bhat says cf. Kurukh *cho*, Skt. *hyaḥ*. The second is not very convincing. Again, Bhat is reaching for it. (M *sé*.)
44. *cheri* (B) ‘goat’. Bhat says cf. Korku *siṛi*, Skt. *chāgala*, Bengali *cheli*. Amen to it, especially Bengali. (M *seri*.)
45. *chikār* (B) ‘hunt’. Bhat says cf. Hindi *sikār*. (M *sikāri* ‘hunter’.)
46. *chokra*, *sokṛa* (B) ‘bread’. Bhat says cf. Korku *sokra*. Amen to it. (M *sokorā*.)
47. *chunḍuku* (B) ‘box’ (M *sandako*).
48. *dada* (K) ElBr. Wow! Talk about child speak. (M MoBr, Fa.)
49. *dbankar* (K) ‘shepherd’, ergo ‘herder’. (Misprint for *dhankar*?)
50. *dēwta* (K) ‘god’. Cf. Indic and Dravidian from Indic, back to *deva*. Specifically, Marathi *devta* ‘god’, but *dev* ‘sun’ says Asha Mundlay. (B *ḍevta* ‘sun’; M *dewtā* ‘sun’.)
51. *diya dēwta* (K) ‘sun’. Maybe means ‘god of day’? No, means ‘light of god’, says Asha Mundlay.
52. *dhatta* (K) ‘cow’.
53. *ḍanḍo* (B) ‘upper arm’ (cf. M *dānḍo* ‘handle’?).
54. *ḍia/dia* (B) ‘day’. Bhat says cf. Skt. *divā* (and Spanish!).
55. *ḍin-oka* (B) ‘daily’. Bhat says cf. Hindi for *din*.
56. *ḍoŋgor* (B) ‘forest’ (cf. M. *ḍongor bérko* ‘wild cat’).
57. *ḍuḍ/dud* (B) ‘milk’. Bhat says cf. Bengali et al., Skt. *dugdha-* (M *dudo*).
58. *ḍukri may* (B) FaSi (M ‘grandmother’, FaElSi).
59. *eḍ-* (B) ‘to come’, *ēdē* (K) ‘to go’ (M *én* ‘to come, approach’).
60. *ējē* (K) ‘bad’.
61. *er-/iyr-* (B) ‘to go’. Bhat says cf. Parengi *iai-*, Sora *iy-*, *yir*. Amen to it. (M *ér* ‘to go’, *iyér* ‘to come’).
62. *etarēn* (K) ‘his, their’. Third person pronoun, singular or plural, possessive. (M *eté* ‘he’.)
63. *ēthē/ethe* (K) ‘to be, was’ (third person only).
64. *gā* (K) verb suffix, present tense, as in *jo koṭṭe-gā* ‘I beat’.
65. *gon* (B) ‘with’ (associative, comitative).
66. *gora kelli* (B) ‘male calf’, ergo ‘bull calf’; *kelli* as such seems to be ‘female’ (M *gorha* ‘young he-calf’, *kelli* ‘she-calf’).
67. *Gullu* name of B’s informant (Gullu Patel of Sonballi village, according to Mundlay).
68. *ghaṭa-* (B) ‘to search’. Bhat says cf. Korku.
69. *ghūrka ed-* (B) ‘to go defecate’, where *ed-* is the ‘go’ part (M *ghúr* ‘anus; to excrete’).
70. *hā* (K) ‘alas!’ Interjection. Contrast with *hā̃* ‘yes’.
71. *haran* (K) ‘deer’, cf. Gadba *harna*, etc.
72. *heŋgen* (B) ‘me’ (first person singular pronoun, object).
73. *heron* (K) ‘duck, sp. bird’.
74. *heṭṭi*, *eṭṭhi* (B) ‘elephant’ (Skt. *hastin-* ‘handed one’, old translation of autochthonous word; cf. Kipling’s “hathis piling teak”).
75. *hey beṭela* (B) ‘there is not’.
76. *hi*, *i* (B) ‘this’ (demonstrative).
77. *hiṅgan* (K) ‘we, our’. First person plural pronoun, focal and possessive. Sono dubbio. (M *ingin* ‘we’.)
78. *hiṅgē* (K) ‘mine’. First person singular pronoun, possessive.
79. *hiṅgē-ṭhāku* (K) ‘of me, of us’. First person pronoun, both singular and plural, possessive.

80. *hō* (K) 'he, they'. Third person pronoun, singular or plural, focal. Note no marker for feminine, apparently, in Nihali pronouns. (M *ho* 'he'.).

81. *howṭa/how-ṭa* (B) 'they/they' (with presumed segmentation: third person plural pronoun, focal). Seems to be 'he plural'. (M *howeṭlā*.)

82. *hoytarē-ṭhākun/hoytarē-ṭhakun* (K) 'of him, of them'. Third person pronoun, singular or plural, possessive. I doubt that the small difference in vowel length as the difference between singular and plural. Methinks they are the same form. Segmentation might be **ho-itarē-ṭhāku-n*.

83. *iŋga* (B) 'here'. Bhat says cf. Korku. (M *hiŋki, hingā*.)

84. *iphil-ta* (K) 'star'. Probably 'stars', i.e., plural. Cf. Munda *ipil*.

85. *itiki* (B) 'here'. May be simple mishearing of **hitiki*. (M *hiṭini, itkā*.)

86. *iṭṭel* (B) 'they two' (third person dual pronoun, focal). (M *iṭikel*.)

87. *jākoṭo* (K) 'male', inferred from *jākoṭo māu* 'stallion', *jākoṭo haran* 'stag' = 'buck'. Cf. Ongota *soqta* 'male (animal)', says Hal.

88. *jappo* (B) 'water' (M *joppo*).

89. *jari* (B) 'root'. Bhat says cf. Korku, Hindi, etc. (M *jaḍi*.)

90. *jere-* (B) 'to remain', ergo 'to be left over' (M *jeré* 'to become, be ready'; auxiliary verb of importance; K gerund of past perfect action).

91. *jilŋguij'* (B) 'earthworm'. [j'], whose phonetic value I do not know ["checked consonant" according to Bhattacharya, p. 246; cf. *kaplij'*, below. Ed.] (M *jilgutiñ*.)

92. *jopaṭke* (K) 'if' (conjunction).

93. *jopo* (K) 'water' (see *jappo*, above).

94. *jūḍ* (B) 'bamboo'. Bhat says cf. Korku *mad*, et al. Non credo. (M *jodu*.)

95. *kā/kē* (K) 'to be' (present tense) / 'to be' (present tense, third person only).

96. *kādini* (K) gerund suffix on main verb root, as in *koṭṭo-kādini* = 'beat-ing'.

97. *kaka* (B) MoYoSiHu, FaSiHu, FaYoBr (M 'uncle').

98. *kaki* (B) FaYoBrWi (M 'aunt').

99. *kakheyŋ* (B) 'to comb hair' (M *kākhén* 'comb' [noun]).

100. *kāmo* (B) 'work', also 'to do' and the causative marker. Bhat says cf. Hindi *kām* (Skt. *karma-*).

101. *kanḍe* (B) 'tuber'. Bhat says cf. Korku; Hindi *kāndā*. Amen to it. (Cf. M *kānde* 'onions'.)

102. *kānṭi* (B) 'for the purpose of'.

103. *kaplij'* (B) 'butterfly', but with no idea what the ['] does to the [j]. [Cf. *jilŋguij'*, above. Ed.] (M *kāpliñ*.)

104. *kaṭan-* (B) 'to be silent'. Bhat says cf. Korku. (M *kaṭām* 'silence'.)

105. *kaṭṭo-/koṭṭo-* (B) 'to beat'. See 'to pound'. Bhat says cf. Munda *kuṭao*, Dravidian *koṭṭu*. Since the similarity extends from Nihali to Munda to Dravidian, it seems this root has probably been borrowed twice. Who had it first?

106. *ke/ki/-ke/-ki* (K) 'in, to' (noun suffix, locative). Third and fourth forms are as shown in citation forms.

107. *kē/ku/kā/-kē/-ku/-kā* (K) 'of' (noun suffix, genitive). Last three forms are as in citations. I have to mention that this like Omotic genitive in *ko, ku*.

108. *kirsan* (K) 'cultivator', ergo 'farmer'? Nihals either do not farm, or they are late in the practice of it. (? M *kirsā* 'rich man'.)

109. *kobḍur* (B): no gloss; apparently = M *kobdur* 'pigeon'.

110. *kon/kom* (B) 'from' (see *kun*, below).

111. *kuguchi* (K) 'hair'. Maybe cf. Malto *kuku* 'head', which is rare in Dravidian, but Kurukh *kukk* 'head, extremity' also occurs. Just as likely to go from Nihali to Dravidian. (B *kuguso, kuguchyo*; M *kugusu, kuguso*.)

112. *kun/-kun* (K) 'from, to' (noun suffix, locative); second is citation. Cf. Hindi *se*, says Grierson (really Konow).

113. *khara* (B) 'field' (M 'earth').

114. *kharuka* (B) 'many' (animate) (cf. M. *khārú* 'family, herd'?).

115. *khuri* (B) 'leg'; (K) 'foot'. Bhat says cf. Skt. *khura* 'hoof', Dravidian *kal* 'leg'. Not quite so convincing. (M 'sole of foot'.)

116. *lānā* (K) 'child, son'. (Apparently not known to B and M.)

117. *laṅg* (K) 'tongue' (see *lãy*).

118. *laŋka* (B) 'a god worshipped in the month of Phalguna (February-March)'. The Korku worship *Rāvaṇa*, who, according to Hindu mythology, is the king of *Laŋka*. Bhat says cf. Korku, where *laŋka* means 'distant place'. (M *laŋkā* 'far-away place; world after death'.)

119. *lãy* (B) 'tongue'. Bhat says cf. Korku *lãŋ* and similar, but so is Romance *lingua*, French *langue* [lãg]. (M *laiñ.*)

120. *lēgē* (K) 'up'.

121. *maikko* (B) 'bee' (M *mãiko(t)* 'mosquito, fly').

122. *mami* (B) MoBrWi. Note: it contrasts with *mama* = MoBr.

123. *mancho/manṭa* (B) 'man/men'. Bhat says cf. Skt. *manuṣya*. This makes sense for the first form, but not the second, which is *man* + *ṭa* = 'man' + plural. (M *maccho mānso* 'healthy man'; *mānṭa* 'neighbors'.)

124. *mancko* (K) 'man'. By inference, the *-ko* is a suffix, perhaps ancient. (Cf. *mancho*, above.)

125. *māṭo* (B) 'thigh'. (M *máʔaṭo* 'leg, thigh'.)

126. *miŋgay* (B) 'where?' (M *mingā*).

127. *minjar* (B) 'palm (of hand)'. Possibly 'sole (of foot)' too. (M *minjārā* 'midst'.)

128. *mindi dēwta* (K) 'moon'. Means 'god of light'. (Cf. *dēwta*, above.)

129. *n/ne/-n/-ne* (K) 'of' (noun suffix, genitive). Third and fourth forms are as in citations. I should mention that an *n* is found in Basque and Caucasic and/or Burushaski as well as Afrasian.

130. *n/ne/-n/-ne* (K) 'by, with' (noun suffix, agentive). This looks a great deal like the genitive in the same form. Mundlay has *-nā* 'by' and 'possession'.

131. *nāko* (B) 'you two' (second person pronoun, dual, focal). Since there is some reason to believe that this segments to *nā* + *ko*, then, see both *nē* 'thou' and *lā* 'you' (pl.). (M *nyeko.*)

132. *nakko* (B) 'nail of finger', ergo 'claw'. Bhat says cf. Skt. *nakha*. (M *nakkho.*)

133. *nāku* (K) 'you, of you' (see *nāko*, above).

134. *nāŋ/nān* (B) 'what?' (M *nānā*).

135. *nānko* (K) 'what?'. Cf. Beja and Ongota, Hal says. (M *nānkā* 'of what sort?')

136. *nāpyom* (B) HuElSi, WiElBrWi (cf. M *nāpyam* and *napyom*).

137. *nē* (B) 'thou' (second person singular pronoun, focal). (M *nyé.*)

138. *nē* (K) 'thou, thine; your (pl.)' (second person pronoun, focal, male or female, both singular and plural, Konow seems to say). Sono dubbio — deeply. It gets more convincing when you assume that 'your (pl.)' is a mistake for 'thy'.

139. *ne-ṭhāku* (K) 'of thee'.

140. *o/jo o* (K) 'to be, was' (first and second persons) / 'I was'.

141. *oḍov* (B) 'buffalo'. (M *odow* 'she buffalo'.)

142. *om/-om* (B) noun suffix, combining form in kinship terms, possibly meaning a senior sibling type affinity; e.g., *nāpy-om* HuElSi.

143. *pachlā* (K) 'behind' (M *pasalā* 'second, other').

144. *pakin* (B) 'peacock' (M 'hen').

145. *pakoṭo* (B) 'bone'.

146. *palco* (B) 'son'.

147. *parayn* (B) 'river' (M *parãy*).

148. *parka* (B) 'all' (cf. M *pár* 'everything').

149. *paroŋ* (B) 'bank of a river'. Bhat says cf. Skt. *pāram*.

150. *pasi-ki* (B) 'near'. Why the segmentation as it is? Bhat says cf. Hindi *pās*.

151. *peṅg* (K) 'head' (see *pẽy*, below).

152. *perijo* (K) 'daughter' (see *pirju*, below).

153. *pẽy* (B) 'head'. The vowel is long and nasalized. (M *pyeŋ.*)

154. *pi-* (*pa-*) (B), *pīya* (K) 'to come'.

155. *pin* (K) 'but' (conjunction).

156. *pirju* (B) 'daughter' (M *pírjo* 'daughter, girl').

157. *poya-ta* (K) 'bird'; *poyye* (B) 'bird'. Methinks *poya-ta* a mistake for 'birds'. See comparative evidence. Maybe Dravidian related, K says. Sono dubbio. (M *poy* 'bird', *poye* 'feather'.)

158. *phuphu* (B) FaYoSi. Bhat says it is from Hindi.

159. *rabanka* (B) 'cold'. Bhat says cf. Korku. If it is highly similar and quite long like this, then the odds of borrowing increase. (M *rābān.*)

160. *raṇḍa* 'boy', as in *bidi ējē raṇḍa* 'one bad boy'. (? Cf. M *rānḍ* 'prostitute'.)
161. *sanu* (K) YoBr. Cf. lots of lingos in the West.
162. *soso* (B) Bhat directs readers to "see *chocho*", which is not listed; M also lists 1507 *soso* with no gloss — does anyone know?
163. *ṭarsya* (B) 'sp. animal, called *taṛas* in Marathi, etc.' Is that clear now? (M *tarsyā* 'hyena'.)
164. *ṭembriya* (B) 'tiger'.
165. *ṭẽya* (B) WiBr, WiSi, ergo 'sibling or wife'. (M 'brother-in-law'.)
166. *thenḍey* (B) 'moon' (misprint for *ṭhenḍey*?: M *thénded*).
167. *ṭhuk-* (B) 'to spit' (Hindi *thūk-nā*) (M *thuk*).
168. *ūcā* (K) 'high' (M *ucā* 'cliff, tall').
169. *ugaen-* (B) 'to remain, to live'. A bit odd in semantics.
170. *uṇṭuṭa* (K) 'camel', cf. Korku *uṇṭo*, K says. It may very well be *uṇṭo* + *ṭa* = 'camels'.
171. *uri-* (B) 'to kindle' (M *ú*, *uru*).
172. *vorcho* (B) 'year' (Skt. *varṣa*).
173. *yēpṭa/yēp-ṭa* (B) 'honey' (with some analysis). Judged as a mass noun, then 'honey' can be plural in form, and *-ṭa* marks the plural. (M *épṭa* 'bees'.)

Nihali Lexicon: Supplement II
Nihali Numerals

Numeral	Konow	Bhattacharya	Mundlay
1.	bidi	biḍi (f, n) biḍum (m)	biḍi, biḍik, biḍiko, beḍe, baḍā
2.	irā	irar (m) ir (f, n)	irār, ír(i)
3.	moṭho	mōṭh(o)	moṭh
4.	nālo	nālo, nālku	nālku
5.	pāñco	pā̃co	pānc
6.	chāh	-	sa
7.	sato		
8.	aṭho		
9.	naw		
10.	das		

As pointed out by Robert Shafer (1940, quoted in Blažek's article in this issue), only the numeral 'one' seems to be native to Nihali. The numerals 'two, three, four' were acquired from Dravidian neighbors at an early stage, and the numerals from 'five' on up from the Indo-Aryans. (Bhattacharya and Mundlay do not list numeral words above 'five, six', presumably because they are all identical with neighboring Indo-Aryan words.)

Note some possibly native ordinals: Mundlay's [1531] *chebela* 'third', [1565] *talāri(re)* 'third', [1566] *ṭalpono* 'fourth', [369] *cayni* 'first', [379] *céyni* 'first, previously'.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali Lexicon: Supplement III
Dental and Retroflex Variants

In comparing the different Nihali glossaries, we have noted a number of cases where words in the dialect recorded by Mundlay have a dental consonant, while Bhattacharya (or Konow) record a retroflex consonant, or vice versa. Since two of these field workers are native Indians, it seems unlikely that all these cases can be attributed to mishearings. Rather, at least some of them must be genuine dialectal variants. We have noted the following instances of this variation. Mundlay's form is cited first, followed by Bhattacharya's (B) and/or Konow's (K).

In a letter to the Editor dated 2 November 1996, Dr. Mundlay states that "*cerebrals* ... are very frequent in Nihali, and whenever you are in doubt, assume a cerebral [i.e., retroflex — Ed.] ... it is possible that I missed some of them in typing." In the same letter, Mundlay clarifies the forms (92) *áṭo* 'tears' and (1580) *ṭéko* 'we' (dual), with retroflex.

Number	Mundlay	Bhattacharya/Konow
87.	*asudi*	(B) *achuḍ-*
89.	*átho*	(B) *ā̃ṭho*
92.	*ãto*	(B) *yā̃ṭo*
114.	*badrā*	(B) *badṛa*
156.	*bardo*	(B) *barḍo*
168.	*batāme*	(B) *baṭam-*
170.	*bātuko*	(B) *baṭuko*
205.	*béthe*	(B) *beṭe*
248.	*botor*	(B) *boṭor*
362.	*cáto*	(B) *cāṭo*
446.	*daḍi*	(B) *ḍaḍi*
448.	*dai*	(B) *ḍai, ḍay*
477.	*deḍḍa*	(B) *ḍeḍḍa*
479.	*delen*	(B) *ḍelen-*
483.	*dewtā*	(B) *ḍevta*, (K) *dēwta*
486.	*din*	(B) *ḍin, din*
494.	*ḍo*	(B) *ḍo*, (K) *do*
529.	*dhāpri*	(B) *ḍhapri*
559.	*dhullā*	(B) *ḍhulla*
568.	*édugo*	(B) *eḍŭgo*
586.	*eté*	(B) *eṭe, eṭey*
589.	*gādáw*	(B) *gaḍao-*
591.	*gadri*	(B) *gaḍri*
602.	*garā*	(B) *gāṛa*
604.	*gardan*	(B) *garḍan*
642.	*gothi*	(B) *goṭhi*
682.	*hardo*	(B) *harḍo*
714.	*hondar*	(B) *honḍar*
758.	*jaḍi*	(B) *jari*
813.	*jodu*	(B) *jūḍ*

Number	Mundlay	Bhattacharya/Konow
868.	*kānde*	(B) *kanḍe*
927.	*kobdur*	(B) *kobḍur*
975.	*khāndá*	(B) *khanḍa*
986.	*kheḍā*	(B) *kheda-*
1053.	*lokhando*	(K) *lokhaṇḍo*
1077.	*mándo/u*	(B) *mānḍo/u*
1106.	*méndi*	(B) *minḍi*, (K) *mindi*
1158.	*murkitij*	(B) *murkiṭij'*
1176.	*nārā*	(B) *nāṛa*
1199.	*nidir*	(B) *niḍir*
1205.	*niṭṭo*	(B) *nitto*
1212.	*odow*	(B) *oḍov*
1282.	*pātāri*	(B)(K) *paṭar-*
1311.	*péte*	(B) *peṭe-*, (K) *pēṭē*
1450.	*sandako*	(B) *chunḍuku*
1570.	*tāndur*	(B) *tanḍur*
1573.	*táro*	(B) *ṭār-*
1577.	*tarsyā*	(B) *ṭarsya*
1580.	*téko*	(B) *ṭēku*, *ṭyēko*
1581.	*té*	(B) *ṭiye-*, *ṭē-*, (K) *ṭē-*
1588.	*thuk*	(B) *ṭhuk-*
1610.	*toṭā*	(B) *ṭōṭa*
1629.	*thénded*	(B) *thenḍey* (misprint for *ṭhenḍey*?)
1654.	*urā*	(B) *uṛa-*

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Much Ado About Nothing

Paul K. Benedict
104 River Lane, Ormond Beach, FL 32176 USA

(Editor's Note: Paul Benedict is an internationally-known and renowned pioneer long ranger. His revised taxonomy of Southeast Asia, detaching the Thai or Daic group of languages from Chinese and the Sino-Tibetan phylum, and creating a new phylum called Thai-Kadai (now simply *Kadai*), was published more than a half century ago. It has been highly influential in studies of Southeast Asian prehistory. His later work connecting Kadai to the vast Austronesian phylum under the rubric of *Austro-Thai* has extended his influence on prehistoric studies to Oceania. Not unexpectedly for a long range hypothesis, his Austro-Thai has met much more resistance than his original Kadai thesis. Most recently, he has boldly asserted the genetic ties between Japanese (and the Ryukyuan group) and Austro-Thai. His latest conception of that has Austro-Thai as a super-stratum on a probable Altaic sub-stratum. In linguistic terms, that essentially means that the main Japanese-Ryukyu group is genetically Austro-Thai but has many old Altaic borrowings. Given the greater popularity of the Japanese-is-Altaic thesis, plus the embattled nature of Altaic itself, Benedict's "Austro-Japanese" has met severe criticism. Still, some long rangers have accepted it, including some Nostraticists. Differing from some long rangers, however, Paul Benedict does not accept the *Austric hypothesis* anymore because of old borrowings and areal influences. The two groups are much interdigitated in Southeast Asia.)

(His remarks have been lightly edited and given a title. HF)

Concerning the classification of Nihali. I quickly ran the Nihali material written by Asha Mundlay (plus Grierson and Bhattacharya) by my "mental tapes" of pertinent data and reconstructions. My "mental tapes" cover, of course, three enormous phyla in Southeast Asia and Oceania. And they work speedier than computers. Yet my "mental tapes" came up with *almost nothing*. Even by chance, we should do better than that, it would seem! What does this mean? Should we assign a smaller role to chance? Do you know of any parallels? If so, this is great news for long rangers. It should be discussed in *Mother Tongue*[1].

My lone prize in the search was Mundlay's #834 *kago* 'taste bitter' which I might relate to Proto-Tibeto-Burman **ka* 'bitter'. Also Mundlay's #835 *kaggo* 'mouth' might relate to Proto-Tibeto-Burman **m-ka* 'opening, mouth, door'. We might set up a disyllabic Sino-Nihali, with *regular* loss of 2nd syllable (in Sino-Tibetan). Great! We need more!

---

[1] We are happy to report the informal but foreseen results of the first major test of a "splitter" theory, viz., how easy it is to find lexical similarities and how worthless they are in genetic classification. (We know this theory best in the vernacular as "You can throw cow dung at a barn and some of it *will* stick".)

Clearly taxonomists do not simply look at "similarities" — which have never been defined properly. Or "look-alikes" truly are not so common. Repeatedly, our referees reported that they could find very little or nothing, i.e., very few similarities that they would call cognates, other than obvious loan words. Reminding the reader that the scope of this inquiry ranged over hundreds of languages, we conclude that the splitters might actually be wrong! [Editors]

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali and Ainu

John D. Bengtson
Minneapolis, MN

The embryo of this essay was stimulated several years ago by F. B. J. Kuiper's (1962:43-44, _et passim._) notes about certain Nihali (and Munda) words that resembled Ainu words. Kuiper hesitated to draw any conclusions from these parallels, since they seemed to be few, and impossible to distinguish from chance resemblances.

Over the past few years, I have collected several other Nihali-Ainu comparisons. Recently, we have had the additional advantages of Asha Mundlay's Nihali lexicon, and on the other side, Alexander Vovin's (1993) reconstruction of Proto-Ainu. (Ainu reconstructions cited below, except in §2, are by Vovin, other Ainu words, except §2 are taken from Batchelor's dictionary. In Vovin's reconstructions, _E_ = open _e_, _O_ = open _o_, = = morpheme boundary, (=) = possible morpheme boundary.)

The expanded list of Nihali-Ainu comparisons now includes:

	Nihali	**Ainu**
1.	M 65, 66 _ápó_ 'wood, to be lit'; _ápokama_ 'to light a fire'; _apóka_ 'it burns'; Bh 24 _āpo_ 'fire'.	*_apOy_ 'fire': Hokkaido _ape_, _abe_; Kuril _apoi_ 'hearth'.
	§Cf. MK: Pear _puy_ 'tinder'; Khasi _dpey_ 'hearth, ashes'; Kadai: Kam-Sui *_puy_ 'fire', Northern Tai *_vii_; AN *_hapuy_ (Lopez) = *x_1_apuy_ (Dyen) 'fire'; Malay _api_, Tonga _afi_, Hawaiian _ahi_, etc. (LP).	
2.	M 89 _átho_ 'husband'; Bh 27 _ā̃ṭho_ id.	_enčiu_ (< *_ent-_) 'man (in prayer and folklore)' (Piłsudski 1912).
	§No parallels in other Austric languages? Cf. also nos. 13, 14, 15, 21, 23.	
3.	M 108, 214; Bh 292 _baḍā_ 'one' ~ _baḍa_ ~ _beḍe_ ~ _biḍi_ ~ _biḍik_ ~ _biḍiko_	*_patEk_ 'only'.
	§Most apparent Austric cognates have _m-_: Santali _miɗ_ '1', Kharia _muḍu_ 'alone, only, single'; Khmer _mûoy_ '1'; but Pareng _boi_, etc. (cf. Bantu: Swahili _moja_ '1', etc.!).	
4.	M 212 _bicá_ 'to ask a question', _bica_ 'why', _bica-_ 'to ask' (LSI).	*_pisi_ 'to ask'.
	§Cf. AN: Philippine *_bicara_ 'to talk'.	
5.	M 319, 324 _cačuko_ 'hot', _čačakāmā_ 'to heat', Bh 142 _cacŭko_ 'hot'.	*_sEEsEk_ '(to be, to grow) hot'.
	§Cf. AN *_segseg_ 'to burn'; Eastern Oceanic *_saka_ 'hot'.	
6.	M 369 _cayni_ 'first'; Bh _ceyni_, M 379 _céyni_ 'first, previously', _caini_ 'before' (LSI).	*_si=nE=_ (*_gi=nE=_ / *_hi=nE=_) 'one'.
	§Cf. Munda *_sɛng_ 'first, before': Kharia _sɛng_ '(go) first', Remo _sənsañ_ 'ahead, in front'; MK: Pacoh _xang_ [sa:ŋ] 'first' (Hayes); Khasi _ši_ '1'; Aslian: Besisi Malacca _ciang_ 'first, in front'; Semang Paya _kaʔ-seng_	

'first, formerly'; Kadai: Li **tsi̯* '1', Gelao *si*; AN **ʔitsᵢa* / **atsa* '1', Tsouic **cani*, Philippine **sa/ŋa* '1' (LP).

7. *eŋ(g)-* 'me/my' (1st person sg. obl.; see Pinnow 1966a:496).
   en 'me' (Batchelor) = **an* (Vovin).

§Cf. Munda **iñ* / **ing* 'I'; MK: Khmer *anh*, Mon *ai*, Bahnaric **ʔiñ*; Aslian: Semai (I) *ʔing*, (II) *ʔeñ*; AN **-ŋku* 'my. Mine' (LP).

8. M 767 *jālú* 'to climb down (from a tree, etc.)', Bh 185 *jalū-* 'to descend'.
   esoro (*e-soro*) 'to descend'.

§Cf. AN: Oceanic **suru(p)*, **soro(p)*, **solo(p)*, Polynesian **solo* 'to descend, go down'.

9. M 787 *jeré* 'to become, to be ready', Bh 194 *jere-* 'to remain'.
   siro-ma 'to abide, dwell', *siro-kka* 'to make motionless'.

§? Cf. AN **DiRih* 'to stand'.

10. Bh 188 *jiki*, M 788 *jiki(r)* 'eye'.
    sik* (gik*/**hik*) 'eye'.

§Cf. Munda: Juang *je-* '*eye' in *je-tej* 'eye-sand' (Mundlay); Santali *jhiki* (in *jhiki miki* 'splendid; to glitter', cited by Kuiper).

11. **-kap-* in M 790, Bh 189 *jiki kāpri* 'eyebrow' (*-kap* + *-ri* [personalizer]; cf. §21, below) (originally 'eye-fur'?).
    **kAp* 'skin, fur' in *shik-kap* 'eye-lid' (= 'eye-skin').

§Cf. MK: Stieng *kup, kuo:p* 'skin, bark'; Aslian: Sakai (Sungai Raya) *tšĕ-kop* 'bark, skin', Jakun *tšun-kop* 'skin'; AN: Polynesian **kape* 'eyebrow(s)'; South Formosan **kaba* 'skin': Kanakanabu *káva*, etc. (LP).

12. M 901 *kāw* 'flesh', Bh 84 *kāv* id. (< **kãw* < **kam* ?).
    **kam* 'flesh, meat'.

§Cf. Kadai: Li **xaam* 'flesh': White Sand Loi *kham*, Shaved Head Loi *ham*, Small Cloth Loi *gom*, etc. (LP). The phonetic change *m* > *w/v* seems to be areal: cf., e.g., Romani *nav* 'name', Hindi (tadbhava) *nāũ*, (tatsama) *nām* < Old Indic *nāma-*; and cf. alternations in Dravidian, e.g., Kannada *tamaṅga* ~ *tava(ṅ)ga* 'platform', Kannada *kavuṅkuṟ* 'armpit' vs. Tamil *kamukkaṭṭu* id. (DEDR 1234, 3081).

Alternatively, Hayes (personal comm.) suggests comparison of the Kadai words here with Ainu **kem* 'blood', Miao-Yao **nćhām* 'blood', Mon-Khmer **cəhām* id., etc. (LP).

13. M 916, Bh 97 *kepā* 'louse'.
    kapo 'nits'.

14. M 933, Bh 100 *kokõy* 'ant'.
    **kiki(=)r* 'worm, insect, bug, fly'.

15. M 931 *kógo* 'snake', Bh 109 *kōgo* id.
    okokko (*o-kokko*) 'snake'.

16. M 959, 960 *kugusu*, *kuguso* 'hair', Bh 89 (also) *kuguchyo*.
    kishki /kiski/ 'animal hair'.

§Cf. Munda: Santali *goco* 'beard, mustache', Mundari , Kharia *gucu* id. (Kuiper). The comparison requires either metathesis of the type *KVK(V)SV* ~ *KVS(V)KV*, or differential affixation of the type *KV-KVSV* ~ *KVS(V)-KV*.

17.	M 1079 **manê* 'we' (pl.), Bh 330 *māney* id. (*mVn-* < **bVn-* or **wVn-* ?).	*un* 'us' (obj.) (< **wun* ?).

§Cf. Munda *bi̱(n)* 'we' (incl.); MK: Bahnaric **bi:n* id.; Miao-Yao **(m)pua* 'we'; AN: Philippine **mami* 'we' (excl.), Oceanic **-mami* 'our' (excl.), etc. (LP).

Vovin (personal communication) notes that the Nihali pronoun looks "Nostratic": cf. Indo-European **me-* 'we'; Uralic: Mordvin *miñ* 'we'; Dravidian: Telugu *manamu* 'we' (incl.); Altaic **bän*; Japanese **bàn[u]*, etc.

18.	M 1106 *méndi* 'night', Bh 337 *minḍi* 'evening, night'.	**mOn(=)rE* 'to be late at night'.

§Cf. Munda: Santali *ñindạ* 'night', Mundari, Ho *nida* id. (Kuiper disputes a connection, but it seems possible by assimilation or dissimilation.)

19.	M 1109 *ménge* 'teeth', Bh 341 *menge* 'tooth, jaw'.	**ima(=)k* 'tooth' (Yakuma *mimak*).

§Cf. MK: Mon *ṅeak* 'tooth, jaw', Khmer *thméñ* (*dhméñ*) id.; Aslian: Sakai *lĕmoing*, Semang *lamoing* id. (Kuiper); Kadai: Kam-Sui **hmaak* 'to chew'; AN **mamaq* 'to chew': cf. Oceanic **maka(s)* 'tooth', **maŋa* 'mouth' (LP).

20.	M 1163, 1164 *nānā* 'what', *nāni* 'who', M 1186 *nāwki* 'in what way', M 1188 *nawā* 'of what type'.	**nEE=* 'who, what', *nen* 'who', *nep* 'what'.

§Cf. MK: Katuic **n□w* ~ **naw* ~ **nə-na:w* 'who'; Khmer *ṇa* (*ṇā*, *na*) 'which' (Pinnow); AN: Philippine **nenu* 'what'. (*N*-interrogatives are found in diverse languages, e.g., Swahili *nani* 'who', Japanese *nani* 'what', Basque *no-* [interrogative stem], etc.)

21.	M 1304 *pephryā* 'saliva' (analyzed as *peph-ry-ā*? cf. *-ri* in 'eyebrow', §11, above).	**pOp* 'sweat'.

§For 'saliva' ~ 'sweat', cf. Toda *ko·ẕ* 'saliva, drool': Gondi *kālum* 'sweat' (DEDR 1478).

22.	M 1314, 1315, Bh 268, 269 *petek-* 'to break, to tear'.	**pet=u* 'to cut, to split'.

§Cf. Munda: Santali *pe̱ṭe̱ć* 'to break off with the finger, nip off, snap off', Mundari *pete'* 'to break off a twig or small branch', etc. (Kuiper); AN **bəṭak* 'to split'.

23.	M 1306 *pyeŋ* 'head', Bh 271 *pẽ̄y*, LSI *peṅ* id.	**pa* 'head'.
24.	Bh 256 *paṭ-/piy-* 'to come', Bh 261 *pi-* (*pa-*) 'to come'.	**pay=i* 'to go'.

§Cf. Munda: Gutob *pi-* 'to come'; MK: Bolowen *buh*, Alak *boh* id.; Aslian: Semang Plus *peh*, Sakai Tanjong *bhi*, *bej* 'to come'; Kadai: Kam-Sui **pa:i* 'to go, walk', Tai **pai* 'to go' (LP).

25.	M 1580 *téko* 'we' (dual), Bh 202 *ṭyēko*, *ṭēku* 'we two' = *té-ko*, *ṭyē-ko*, *ṭē-ku* (*-ko*/*-ku* is a separable lexeme: cf. *nā-ko* 'you two').	**ti*= 'we' (pl.) realized as [či] in all dialects.

§This appears to be the Nihali inclusive 'we' as opposed to the exclusive *manê* (no. 17, above); cf. Kadai: Tai **tuu* 'we' (exclusive), Mak *di* 'we' (excl.) / *da* (incl.), Lakkia *ta* / *tau* id.; AN: Western Fijian **ti* 'we' (incl., trial, present/future), **tu* 'we' (incl., trial, non-time/past), etc. (LP).

26.	M 1656 *ú* 'to kindle', M 1660 *uru* id., Bh 251, 253 *uri* id.	**uguy* 'to burn' vi; **uu[y]na* 'ash(es)'; **unti* 'fire'.

§Cf. MK: Bahnaric *ʔuñ* 'fire': Stieng *uiñ*; Katuic **ʔu:jh* 'fire'; Khmer *ʔuh* 'firewood'; Kadai: Tai **viin* 'firewood' (LP).

In general, I think these comparisons are semantically and phonetically precise, and of a basic character. Many of the semantic equations are exact ('hot' ~ 'hot', 'flesh' ~ 'flesh', 'tooth' ~ 'tooth') or nearly exact ('one' ~ 'only', 'louse' ~ 'nit', 'come' ~ 'go'), exemplifying commonplace shifts of meaning.

The phonetic equations are almost always unremarkable, and, even in this small corpus, observably regular: e.g., Ainu *s* corresponds in several cases to Nihali *c* and *j*.

All the examples are also *basic*, representing the least mutable and most stable lexical elements of both languages. These kinds of words are of the greatest value in determining the genetic affinities of languages, especially when the connection may be quite remote. In particular, the meanings 'who/what', 'eye', 'tooth', 'louse', and the pronouns 'me' and 'we/us', all represented above, are generally considered to be among the most stable of all.

If these parallels, or even some of them, do indeed turn out to be real cognates, it would of course not require a hypothesis of direct contact between India and Japan. A model that would explain them is the hypothesis of an *Austric* macrophylum (see recently Pejros 1992, Hayes 1993), a very old linguistic stratum that once dominated virtually all of southern and southeastern Asia, and has given rise to the modern Austroasiatic, Miao-Yao, Austronesian, Kadai, and (if the thesis of this paper is correct) Nihali and Ainu.

The idea of the Austric (or Austronesian, or Austroasiatic) affiliation of Ainu is not new, and was championed long ago by Olof Gjerdman (1926) and Leo Sternberg (1929); and more recently by Vovin (1992, 1993), Shichiro Murayama (1992a), Václav Blažek (see his article in this volume), Paul Sidwell (1996), and the present writer (Bengtson 1992, Bengtson & Blažek 1996).

One will note that, while some of the Nihali-Ainu comparisons (notably [1] 'fire', [8] 'I/me', and [24] 'come/go') have parallels in other Austric languages, some others (2, 13, 14, 15, 21, 23) seem to be restricted to Nihali and Ainu. To the extent these latter are true cognates, they can be interpreted as relic words preserved only in these two outposts of the Austric macro-phylum.

Abbreviations

AN = (Proto-)Austronesian
Bh = Bhattacharya (1957)
DEDR = *Dravidian Etymological Dictionary*, revised (Burrow & Emeneau 1984)
LP = "Lexical Parallels" (Bengtson & Blažek 1996)
LSI = *Linguistic Survey of India* (Grierson 1906)
M = Mundlay's "Nihali Lexicon" (in this volume)
MK = Mon-Khmer

[PS: "Salting" the lexical salad:]

There is also a remarkable resemblance between the respective words for 'salt': Nihali *cópo* and Proto-Ainu **sippO*. Kuiper thinks the first is a loanword from Dravidian (or "Pre-Dravidian"), while Vovin considers the latter a loan from Japanese. Dravidian has, e.g., Parji *cup*, Kolami, Gadba *sup*, Tamil, Telugu *uppu*, pointing to **cuppu*, identical with the Gadba Salur form cited by Kuiper. Vovin derives the Ainu word from Old Japanese *sipo* 'salt' (Modern Tokyo Japanese *shio*). Vovin (personal communication) has also informed me that the latter word, *sipo*, has been compared with Proto-Korean **sòkòm* 'salt'. If this is correct, the original second consonant may have been *k*$(^{w})$, not *p*.

There are similar words in distant parts of the world, e.g., in Amerind: Salvadoran Lenca *¢'epe* 'salt', Honduran Lenca *sepe*; also with related meanings ('sour, bitter'); Japanese *suppai* 'sour'; Finnish *hapan* id., *sappi* 'gall'; Arabic *ṣafrā* id.;

Cf. with *w*/*v* or *m*:

Uralic: Hungarian *só*, *sava-* 'salt'; Dravidian: Gondi *sawwor* 'salt'; Amerind Papago *síw* 'bitter', Algonkian **ši:w-*/ **šawa:-* 'sour, salt'; Na-Dene: Sarsi *zúú-*, *zúw-* 'sour'; Sino-Tibetan **sūr* id.; Hungarian *sav-anyú* id.; English *sour*; Caucasic **c̲'ĕmhV* (~ **c̲'wĕnhv*) 'salt': Avar *c̄'am*; Basque *samin* 'bitter', etc.; Austronesian **hasəm* or **qasəm* 'sour' (Malay *asam*, etc.); Tai **som* 'sour'; Uralic: Zyryene *šom* 'leaven, sourness'; Hebrew *šemer* 'yeast, dregs' (Ainu *šipuš-kerek* 'yeast'), Hebrew *ḥāmūṣ* 'sour', *ḥōmeṣ* 'vinegar'; Egyptian *ḥmˁt* 'salt', Coptic *hmou*; Amerind: Hianocoto **haamɨ̄-* 'salt'.

Under the circumstances, it seems advisable to separate the Nihali and Ainu words for 'salt' from the proposed genetic evidence, though these words are of interest for other reasons.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Seeking the Relatives of Nihali

Václav Blažek
Masaryk University
Příbram / Brno, Czech Republic

The position of Nihali in genetic classification remains obscure in spite of the efforts of several prominent scholars. Let us repeat the most qualified conclusions:

1. "Nahālī is not and probably was never a Muṇḍā language of the same kind as Kūrkū ...

   Despite some apparent correspondences between Nahālī and Tibeto-Burmic, there is no genetic relationship between the two, unless it can be established that there is such a relationship between Austroasian and Tibeto-Burmic. The apparent correspondences are probably accidental.

   While the base of the languages is not Kūrkū, nor even Mundic, that does not mean it is not Austroasian. ...Yet there are many common words in Nahālī which show a close resemblance to corresponding words in these and even Austronesian languages.

   The history of the Nahāls, indicated by the language as we know it, may be surmised as follows. That there was a proto-Nahalian group, judging by the verbs. That the proto-Nahalians came under the dominating influence of the Austroasians, probably most of the vocabulary becoming Austroasian during that period. That subsequently, probably while the Nahāls had a low state of culture, they entered into commercial relations with the Dravidians and adopted from them the numerals from "two" to "four" and a few other words. That finally they came into contact with the Kūrkūs and Indic peoples in Nimar where they now live, adopting many words of all kinds and much of the grammar from one or the other of these dominant groups.

   (Shafer 1940:341-343)

2. "Nahali, a language spoken in Central India and bordering on the Kurku area, has not yet been considered. Its classification is particularly difficult, for the language is not yet sufficiently well known or investigated. It is possible that Nahali is completely separate, as R. Shafer assumes, but it may also constitute a separate branch of Austroasiatic. It is at any rate not Munda. Nahali exhibits a number of words that cannot be explained as Austroasiatic, as Dravidian, or as Indo-Aryan. Its morphological system, on the other hand, is obviously connected with that of the Munda languages: thus, for example, all the Nahali tense suffixes may be compared with corresponding suffixes in Munda. The reliability of these comparisons is, however, reduced through frequent divergences in meaning. The present state of investigation does not permit any definite judgment."

   (Pinnow 1963:151)

Later, Pinnow (1966b:188-191) presents concrete data demonstrating cognates and differences between Nihali and Proto-Munda / Austroasiatic systems of pronouns and verbal morphology. He concludes:

3. "Though almost all of the occurring elements permit comparison with those from the Munda languages — some, due to strongly divergent meanings, with very great reservations — there remains the question of borrowing. The comparisons cited establish parallels chiefly with Kurku, which is indeed a neighbor of Nahali, but to which it stands in no particularly close relationship. Should it be part of Austroasian, comparison would have to be carried out by way of Proto-Munda ...

   For the present, we may state that the verb system of Nahali resembles that of Proto-Munda in all its general features: the lack of incorporated pronouns, presence of a complex of aspect affixes (with great formal similarities), absence of the absolutive, secondary internal transformation in the particles of negation. The postpositions of the aspect (as well as tense) affixes — after the particle of negation — in Nahali, exhibits a rather free and loose position of the aspect affixes. Originally they probably were not bound morphemes (affixes) at all, but rather were independent particles. If it were possible, at a future time, to offer proof for a

> relationship between the aspect affixes of Nahali with those of Proto-Munda — here we could only show a possibility — there would then follow that these aspect affixes of Proto-Munda also go back to independent words or particles; a further step would thus have been taken to bridge the gulf between the analytical Khmer-Nicobar and the synthetic Munda languages."
>
> (Pinnow 1966b:188-191)

Some comments on Nihali (/Munda) and Ainu exclusive parallels, and Nihali words without etymology.

Nihali *carko* 'black-faced monkey' // Munda: Mundari *sara'* 'monkey, baboon', Ho, Kurku *sara*, Dharni (dial.) *saraq* // Ainu *saro* // Japanese *saru* 'monkey' (Kuiper 1962:44, 66). Japanese *saru* is probably related to Written Mongolian *sarmaγčin* id. (Murayama 1966:155). Ainu *saro* looks like a borrowing.

The Munda-Nihali isogloss has a hopeful cognate in Aslian languages: Kenaboi II *sĕró* 'monkey' (SB 660, M 149). Confronting it with the monkey-names: Semang Plus *tarau*, Serting *tĕráu*, Serau, Darat, Jelai *rauh*, etc. (SB 659, M136), the root **(-)rauk* can be isolated.

Munda 'dog' attested in Kherwari, Ho, Asur *seta*, Kurku *sita*, *tsita*, *ćita*, Mowasi *sita*, Santali, Mundari *sītā*, etc. (Bhattacharya 1966:34, #47) was compared with Ainu *seta*, *sita*, etc. 'dog' already by Gjerdman (1926:73-74). The same author has found the "missing link" in one of the Austronesian languages of Formosa: Favorlang *zito* 'een klein hondeken' (Happart 1650, see Gjerdman 1926:51) = *zito* 'dog' (Terrain de Lacouperie, JRAS 1887:489).

Nihali *átho* 'husband', compared with Ainu *enčiu* 'man' (Bengtson #2), has a more hopeful cognate again in Aslian: Mantra of Malacca *thou* 'husband' (SB 637, H180), Sakai of Tanjong Sambutan *taú* 'male', Sakai of the Kerbu river *toh*, (Brooke) *tau*, Tembi Boy *ata-ú* 'husband'; cf. also Southern Nicobar *otō̆hĕ* (*otāha*) 'male' and perhaps Khmer *phdey* 'husband' (SB 652, M16).

Nihali *cago* 'stone' (after Kuiper 1962:65, #94, and 1966:76 isolated) can be compared with the Aslian word for 'hill': Besisi (Hervey) *ch'ogn*, Besisi of Sepang *chong*, Pantar Kapur (Logan) *sēng*, etc. or 'mountain': Serting *ch'hogn*, Sakai of Blanja (Swettenham ms.) *chökn*, etc.; cf. Stieng *sing* 'hill(side)', Central Nicobar *chòng* 'high' (SB 631-632, H90).

Nihali *cigam* 'ear', *cikn-* 'to hear' (after Kuiper 1962:66-67, ##114, 116; 1966:76, 77 unrelated and isolated) resemble Nanhang *chong* 'to listen'. SB 587, E6 compare it with Sedang *töng*, Alak *tong*, Stieng *tang* 'to hear' (Efimov 1990:76 reconstructs Proto-South Bahnar **tsʌ́ŋ*), Aslian: Semang Paya *tĕngi*, Sakai Guai *ting-oi* id., Serau, Darat *gertik*^{n} 'to listen' and Serting, Besisi (Hervey) *tĕgn*, Tembi, Serau, Darat *entak*^{n} 'ear', etc.

Nihali *cópo* 'salt' is really very probably of Dravidian origin (**cup*), and Ainu **sippo* < Japanese *sipo*. There are several different stems among Finno-Ugric words quoted by Bengtson:

1. **čawз* // **čapa* > Finnish *hapan* 'sauer', Mordvin (Erza) *čapamo*, Mordvin (Moksha) *šapama*, Mari *šapə̑* id., Xanty *suw-* 'sauer werden', Mansi *šēw̯-* 'säuern', Hungarian *savanyú* 'sauer' (UEW 54; the connection of Hungarian *só*, *sava-* 'salt' is rejected).
2. **čemз* > Udmurt *šẹm* 'Geschmack', Komi *šom* 'Sauerteig, Säure', Xanty *čim-* 'gären, aufsteigen (der Teig)' (UEW 57).
3. **säppä* > Finnish *sappi*, Lapp *sap'pe*, Mordvin *sepe*, Udmurt *sep*, Komi *sẹp*, Mansi *tā̆p*, Hungarian *epe* 'gall' (UEW 435-436).
4. **šɤwз* > Mari *šua-* 'salzig werden, gären', Udmurt *šu-t-* 'gären machen, säuern', Komi *šu-ź-* 'sauer werden, sauern' (UEW 798).

At most, perhaps only item 4) is compatible with Dravidian and Old Japanese.

There are promising cognates in Afroasiatic too: East Cushitic *_ʕašub-_ 'salt' > Afar _ʕasbo_, Somali _ʕusbo_, Boni _usúúb_^{ə}, Bayso _esebo_ 'salt', Oromo _aššabō_ 'salt in small pieces' (> Gurage _assäbo_, etc.) // Central Cushitic *_səw-_ or *_cəw-_ 'salt' > Bilin, Kemant _šəwa_, Awngi _číwí_, etc. (Gurage _ṣew_, Amharic _čäw_) /// Late Egyptian _sw'b_ 'natron'.

Korean *_sòkòm_ 'salt' is comparable with Udi (Tungus) _sa'aj_ 'salt' (*_saqaj_) and perhaps Ugric *_ćy̌kkɜ_ 'salt' (UEW 839; Mansi _śäχ_ > Selkup _šeäk_ 'salt').

The Egyptian word for 'salt' is correctly _ḥmꜣ.t_ (Middle Kingdom), Coptic (Sahidic and Boheiric) _hmu_, (Fayyumic) _hemu_ = *_ḥamāꜣt_ < √_ḥml_, comparable via metathesis with Arabic _malāḥat_ 'salt taste', derived from _milḥ_ < Semitic * _milḥ-_ 'salt' (W. Vycichl, _Dictionnaire étymologique de la langue copte_. Leuven & Paris: Peeters 1984:299).

Hebrew _šemer_ 'yeast, dregs' is related to Arabic _ṯamūla_^{h} 'residue, remnant, dregs (of a liquid)', derivative of _ṯamala_ 'he supported, aided, protected' (Klein 1987:668). It is evident that the original semantic motivation was different. Similarly, Arabic _ṣafrāʔ_ 'gall, bile', meaning primarily 'yellow' (as in English _gall_).

Tai *_som_ and Austronesian *_h/qasəm_ 'sour' represent probably a common Austro-Tai heritage.

I am not competent to comment on the Amerind examples.

Nihali _kāw_ 'flesh', compared with Ainu *_kam_ 'flesh, meat' (Bengtson #12) is comparable also with Serting (Aslian) _kĕbó'_ 'body' (SB 541, B326).

Nihali _méndi_ 'night', compared with Proto-Ainu *_mOn(-)re_ 'to be late at night', has perhaps a more promising cognate in Pangan of Sam and Galas (Aslian) _mendöi_ 'last night' (SB 572, D19).

Nihali _manê_ 'we', compared with Ainu _un_ 'us' besides Munda *_bi̠(n)_ // Bahnaric *_bi:n_ // Miao-Yao *_(m)pua_ // Philippine *_mami_ 'we', Oceanic *_-mami_ 'our' (Bengtson #18), has an interesting cognate in Besisi of Malacca (Aslian) _'ma_ 'we' (SB 754, W54).

Nihali _miṅgay_ 'where' (Kuiper 1962:91, #374; 1966:78 isolated) has the closest cognates in Aslian: Jakun (Stevens) _ming_ 'where', Jakun of Malacca _mĕnung_ id. (SB 757, W82).

Nihali _parayn_ 'river' (Kuiper 1966:78 isolated) resembles Kenaboi I _par_ 'rain, water'; Long Kiput, Lelak _prar_, Narom _pĕrar_, Dali, Lemeting _perar_ 'rain' (SB 689, R9).

Nihali _pephryā_ 'saliva', compared with Proto-Ainu *_pOp_ 'sweat' (Bengtson #21), resembles perhaps better South Bahnar *_breyyu:_ 'saliva; dew' (Efimov 1990:81, 140).

Nihali _téko_ 'we', compared with Proto-Ainu *_ti-_ 'we' (Bengtson #25), contains the plural marker _-ko_ (as in _nee-ko_ 'you'). The root *_té-_ (_ṯyē-_ after Bhattacharya) is comparable with Nicobar: Nancowry _cüa_, (DeRoepstorff) _tiūe_, _tiĕ_ 'I', Car _cu-ö_ id. (Pinnow 1966b:190). The Austro-Tai cognate 'we' is quoted by Bengtson, 1.c.

Nihali _uri_ 'to kindle (fire)' (after Kuiper 1962:106, #500, isolated) is probably of Dravidian origin, cf. Kannada _uri_ 'to burn, glow, blaze', Telugu _uriyu_ 'to burn, be afflicted, grieve', etc. (DEDR #656).

Conclusions:

1. The number of common Nihali-Ainu isoglosses is too low to deduce any closer relationship or contact between their ancestors.
2. The most convincing comparisons can almost always be supplemented by comparanda from various Austric languages. My purpose was to demonstrate the remarkable closeness of Nihali-Aslian parallels.

3. The most natural solution is to assume the status of independent branches within Austric for both Nihali and Ainu. They may represent the relics of the first waves in western and eastern directions respectively, following upon the disintegration of the common Austric dialect continuum.

Abbreviations:

DEDR = Burrow & Emeneau 1984
SB = Skeat & Blagden 1906
UEW = Rédei 1988

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali and Nostratic

Aaron B. Dolgopolsky
University of Haifa

The Nihali data, sent to me by Hal Fleming, reached me on 2 January 1997. Since I was asked to send my opinion urgently, I had to study the matter for two days only. Therefore, this is nothing more than a quick glance of a Nostraticist at the Nihali lexical stock (and elements of grammar). I had no chance to study alternative connections, especially those of Nihali with Munda and other Austroasiatic languages, which is indispensable for definite conclusions.

Even a quick and superficial comparison of Nihali with the Nostratic lexical material of my *Nostratic Dictionary* (in preparation) suggests that Nihali has some words comparable with Nostratic. What is most important are the personal pronouns and certain morphological elements shared by Nihali and the known Nostratic languages.

Now follow the lexical (and grammatical) resemblances that I have managed to find.

1. Personal Pronouns

The 1 person singular pronoun in Nihali is *jo* (the symbol [j] represents a voiced palatal stop; on Nihali *jo* and the consonant in question, cf. Pinnow 1959:28, 186). The word for 'thou' is *nee* (according to Kuiper) or *ńe* (according to Mundlay). The system of pronouns *jo* ~ *ńe* reminds one of Proto-Dravidian **yaan̲*/**yan̲-* ~ **nii(ŋ)*/**nin-* (according to Zvelebil). The Dravidian system is an innovation within Nostratic, explained by a series of analogical changes (see Dolgopolsky 1984:83, 100). If Nihali does belong to Nostratic, it shares this important innovation with Dravidian, which suggests that Nihali and Dravidian form a special branch of Nostratic. On the other hand, these pronouns are totally different from those existing in Munda and other Austroasiatic languages (see Pinnow 1959:71, 186, 374). The only person-and-number marker reminding one of Munda and Austroasiatic is *iŋ* ~ *iŋgee* ~ *eeŋgee* ~ *hiŋgee* 'my' (cf. Santali *ińˊ*, Mundari *aiŋ*, *iŋ*, Khasi *iŋ* ~ *ińˊ* 'I' [cf. Pinnow 1959:186]), but it may be explained in the framework of Nostratic as well: Nostratic **HoyVnu* 'my' (**Hoya* 'I' + **nu* 'of' > Elamite 1 sg. marker *hu-n* and Dravidian **yaan̲* [cf. Dolgopolsky 1984:100]). Kuiper (1966:74) believes that Nihali *eŋge* 'my' is a loan from North Dravidian, which is hardly imaginable (a borrowed possessive pronoun?). Neither can I accept Kuiper's hypothesis (1966:75) supposing that Nihali *nee* 'thou' is a loan from North Dravidian. Nihali (Mundlay) *mané*, (Kuiper) *maaney* 'we' (pl.) is a typical Nostratic pronoun, reconstructed as **miHanu* 'our' = **mi* 'I' + **Ha* plural + **nu* genitive and found in Chadic (Angas, Karekare *mun* 'we', etc.), in Mongolian *manu* 'our', Dravidian (Gomma Gondi *man̲n̲* 'we' [exclusive], Telugu *man̲əmu* 'we' [inclusive], etc. [cf. Andronov 1978:250-256]) (see Illič-Svityč 1976:52-56). Kuiper (1966:78) includes Nihali *maaney* among words of unidentified origin.

2. Genitive suffix *-n*, *ne*

It is identical with Nostratic **nu* 'of' (postposition) (> Uralic, Turkic, Tungus *-n*; Mongolian *-n* 'of' and *-nu* in *mi-nu* 'my'; Japanese *no* 'of'; Indo-European and Dravidian marker of oblique cases, etc. [see Dolgopolsky 1984:92; Illič-Svityč 1976:79-81; Menges 1960]), as well as with the genitive **-n* in Caucasic, Burushaski, etc.

3. Locative-directive *-ke* ~ *-ki*

It reminds one of Nostratic **k'V* 'to' (cf. Illič-Svityč 1971:368-369), which is found in Afroasiatic, Uralic, Indo-European, and Dravidian.

4. Nihali *khuri*

Nihali (Grierson, Bhattacharya) *khuri* 'foot' may be compared with Nostratic **k'urV* 'animal's foot, hoof' > Tungus **xuru-n* 'hoof' (> Nanay *xorõ*, *xoroã*, Orok *xuru*, etc.), Dravidian **kurVcV* 'hoof' (> Tamil *kuracu*, Kannada *gorasu*, Telugu *gorija* 'hoof' [see Burrow-Emeneau 1984:#1770]), in the meaning 'hand' in Uralic (**kurmV* 'handful'), Mongolian, etc., as well as in a Nostratic phrase **k'urV śulV* 'heel of the foot' (**k'urV* 'foot' + **śulV* 'heel') > Kartvelian **kursl-* 'heel' (Klimov 1964:200) and Semitic **k'ursull-* 'ankle bone' (> Syriac *k'urs'al-aa*, Hebrew *k'arsol*, etc.). The Dravidian word **kurVcV* may go back to **k'urV śulV*, which will explain the presence of the sibilant **-c-* (< Nostratic **ś*), but a Dravidian simple stem ***kurV* (< Nostratic **k'urV*) could also exist, which is suggested by the Sanskrit loan-word *khura* 'hoof'. Of course, we cannot rule out the possibility that either this Indo-Aryan **khura* (namely, its descendant Marathi *khura*) or Dravidian ***kurV* is the source of Nihali *khuri*. Kuiper (1966:72) believes that the Nihali word is a loan from Kurku.

5. Nihali *čoon*

Nihali (Grierson, Kuiper) *čoon* 'nose' may be connected with Nostratic **č'iwnV* 'smell' (especially 'bad smell, bad-smelling sweat') > Uralic **čiwnV* 'smell' (> Mordvin [Erza dialect] *čiyńe* id., Norwegian Lapp *ciwnâ* 'which has a bad smell'), Semitic **√θ'-w-n* (> Mehri *θ'ooneet*, Harsusi *θ'ənweet* 'bad-smelling sweat'), Kartvelian **č'enč'-* (Georgian [dialectal] *č'enč'-* 'to sweat; sweat'). Kuiper (1966:76) includes the Nihali word among "words of unidentified origin and isolated". But on p. 86, he compares the word with Tibeto-Burman dialects, namely, Vayu *čoʔno* 'nose'. All other Tibeto-Burman parallels are phonetically remote (Rai *unu*, Bahing *neu*, etc.). Might Vayu *čoʔno* be a loan from Nihali or some extinct language of the Nihali family?

6. Nihali *peŋg*

Nihali (according to Grierson) *peŋg*, (according to Bhattacharya) *pẽ:y* 'head' — Nostratic **bäńɡe* (or **bäńʕe*) 'head, brain' (the symbol *ɡ* represents a voiced uvular stop) > Uralic **päŋe* 'head' (> Finnish *pää*, Vogul *päŋ*, *päŋk* 'head'), Semitic **√b-n-ʕ* (Akkadian *bibeenu* 'head'), Turkic **bäŋi* 'brain' (> Old Turkic *bäńi*, etc.). An alternative (to my mind, less probable) etymology was advanced by Kuiper (1966:68), comparing Nihali *pẽ:y* with Tibeto-Burman: Sunwar (Darjeeling) *piiyaa*, Thulung *biiu*, and Bahing *piiya* 'head'.

7. Nihali *sanu*

Nihali *sanu* 'younger brother' has an Afroasiatic parallel: Egyptian *sn* 'brother', Cushitic (Beja *saan* and possibly Central Cushitic **-zan/*zän* 'brother' > Xamir *-zan*, Kwara *zän*, Awngi *sén* 'brother'), Chadic (Tsagu *šən*, Bachama *sona*, Mubi *sin*, Jegu *šin* 'brother', etc.) (see Dolgopolsky 1973:295-296; Orel-Stolbova 1995:462-464).

8. Nihali *kooli*

Nihali (Bhattacharya) *kooli*, (Kuiper) *kool* 'woman, wife' (Kuiper [1966:77], word of unidentified origin), (Mundlay) *kalaaṭ* 'wife' — Nostratic **kälu* 'woman of another moiety' (> 'female relative-in-law, bride') > Semitic **kal-at-* 'daughter-in-law, bride' (Hebrew, Akkadian, etc.), Uralic **kälu* (or **kälew*) 'sister/daughter-in-law', Indo-European **ǵl̥ōw-* id. (> Greek *gálōs*, Latin *glōs*, Russian *zolovka*), Dravidian **kall-* 'female relative-in-law' (> Kurukh *xallii*, Malto *qali*), Turkic **kälin* 'bride', Georgian *kal-* 'woman' (cf. Illič-Svityč 1971:295-296).

9. Nihali _ḍiya, diya_

Nihali (Bhattacharya) _ḍiya, diya_ 'day' (unless a loan from Kurku) — Nostratic *_tiyu_ 'light (lux), daylight' > Kartvelian *_te-_ 'light', Indo-European *_dyew-_ 'daylight'. Kuiper (1966:70) considers the Nihali word to be a loan from Kurku.

10. Nihali _eḍ-_

Nihali (Bhattacharya) _eḍ-_ 'to come' — Nostratic *_ʔat(h)V_ 'to come' > Semitic *√_ʔ-t-y_ or *√_ʔ-t-w_ id. (in Hebrew, Aramaic, Arabic, Ethiopic), Cushitic (Beja _ʔat-_ 'go'), Turkic *_aat-_ 'to step', Dravidian *_aaṭ-_ 'to move' (intr.).

11. Nihali _ṭiye-, ṭee-_

Nihali (Bhattacharya) _ṭiye-, ṭee-_ 'to eat' — Nostratic *_t'u(h)ye_ 'to eat; food' > Semitic *√_t-w-y_ (Mehri √_t-w-y_, Akkadian _ta'û_ 'to eat'), Egyptian _t_ 'bread', Cushitic (Beja _tiyu_ 'nourishment'), Chadic (Hausa _či_, Dera _twi_, Migama _tiyaw_ 'to eat', etc.), Tungus *_tuyu_ 'to offer food to a guest', Dravidian *_tu(y)-_ 'to eat' (Tamil _tu-_, etc.).

12. Nihali _iyeer-_

Nihali (Bhattacharya) _iyeer-_ 'to come' — Nostratic *_ʔeyo_ 'to go, come' > Egyptian _ỉy_ 'to come', Berber *_ʔ-y-w_ id., Cushitic (Beja _ʔi-/ʔe-_ id., etc.), Omotic (according to Blažek) *_yi-_ ~ *_i-_, Chadic *_y-_ id., Indo-European *_ʔey-/_*_ʔi-_ 'to go', Finno-Ugrian *_yü-we-_ or *_yi-we-_ 'to come' (> Old Hungarian _jő_, Vogul _yi-/yuw-/yay-_ 'to come', etc.), Tungus *_ii-_ 'to enter' (Evenki, Nanay _ii-_) (cf. Illič-Svityč 1971:197), Dravidian (Krishnamurti) *_ey-_ 'to arrive, reach' (> Kannada _ī-_, etc.) (cf. Illič-Svityč 1971:265-267).

13. Nihali _kerčhi_

Nihali (Bhattacharya) _kerčhi_ 'to itch, scratch' may be connected with Nostratic *_xarčV_ 'to rub, scratch' (the letter _x_ represents a voiceless uvular fricative = <ch> in German _Buch_) > Semitic *√_ħrθ_ 'to plow', Indo-European (Hittite _ḫarš-_ 'to till [the soil]'), Mongolian _arči-_ 'to wipe/clean by rubbing, to erase'.

14. Nihali _eŋger_

Nihali (Bhattacharya) _eŋger_ 'burning charcoal, fire' (unless a loan from Indo-Aryan) may be compared with Nostratic *_HäŋkU_ 'fire' (_H_ = unspecified laryngeal) > Chadic *_ʔak_^{w}_V_ or *_ʔaku_ 'fire' (> Tsaga _áàkwé_, Ngizim _ákâ_, Birgit _ʔàkù_, etc. [cf. Jungraithmayr-Ibriszimow 1994.II:138-139; Newman 1977:26]), Indo-European *_n̥gni-s_ 'fire' (> Old Indian _agni-_, Latin _ignis_, Lithuanian _ugnis_, Old Church Slavic _ognь_ 'fire'), Uralic *_äŋV_ or *_äŋkV_ 'fire; to burn' (> Cheremis _eŋ_ 'fire', _eŋa-_ 'burn without flame', Zyrian _i̮ń_ 'flame', Ostyak _yäŋləl_ 'roast on fire', Hungarian _ég-_ 'burn'), Dravidian *_aṉal-_ 'fire' (> Tamil _aṉal_ 'fire', etc.) (cf. Illič-Svityč 1971:245-246). Borrowing from Indo-Aryan (Sanskrit _aṅgāra-_, Marathi _aṅgar_ 'fire') is not ruled out either.

15. Nihali *bokki*

Nihali (Bhattacharya) *bokki* 'to tie' — Nostratic **bUk'c'E* or **bUk'ć'E* 'to tie; rope' > Kartvelian (Georgian *bac'k'i* 'thin thread'), Indo-European **bhask-* 'bundle' (Latin *fascis* id., Middle Irish *basc* 'necktie', etc.), Uralic *pükse* 'rope' (Mordvin *piks*, Ostyak *püγəŝ* [γ = fricative *g*; *ŝ* = voiceless lateral fricative]), Mongolian *büči* 'rope, tie', Dravidian **pocc-* (Kui *puj-* 'to pack, make a bundle', etc.). Kuiper (1966:76) includes Nihali *bokki* among words of unidentified origin.

16. Nihali *ãṭho*

Nihali (Bhattacharya) *ãṭho* 'husband' (Kuiper [1966:76]: a word of unidentified origin). A curious resemblance to Greek *ánthrōpos* 'person'. No genetic conclusions about this word.

17. Nihali *parayn*

Nihali (Bhattacharya) *parayn* 'river' (Kuiper [1966:78]: unidentified origin) may be connected with Nostratic **biḥra* or **bixra* 'watercourse, river' (*ḥ* = voiceless epiglottal fricative) > Semitic **baḥ(V)r-* 'watercourse, river' (Arabic *baḥr-* 'sea, large river', etc.), Chadic (Angas *feoor*, Sura *vuoγor* 'river, stream', Mubi *burɔɔlɔ* 'river', etc.), Indo-European **bheru-* (Old Indian *bharu-* 'sea'), Tungus **bira* 'river', Dravidian **piir-* 'flow' (Tamil *piir* 'abundant flow') and/or **pUṛa* 'river' (Malayalam *puṛa*, Kannada *poṛe*, etc.) (cf. Illič-Svityč 1967:369).

18. Nihali *khanḍa*

Nihali (Bhattacharya) *khanḍa* 'shoulder' could have been compared to Nostratic **kendV* (or **kenVcdV*) 'shoulder, joint of a limb' > Semitic **√k-n-s-d* 'top of shoulder' (> Mehri *kənsiid*, etc.), Mongolian *qondulay* 'rump, hip (of a horse)', Tungus **kende(-ke)* 'shoulder-blade' (> Orok *kändä*, Evenki *kändäkää*), Dravidian **keṇṭ-* 'ankle, joint of a limb' (> Tamil *keṇṭay* 'ankle', etc.). But it is much more plausible to consider it a loan from Dravidian (see detail in Burrow-Emeneau 1984:#1946) or (together with Kuiper 1966:72) a loan from Kurku.

19. Nihali *čergo-*

Nihali (Mundlay, Kuiper) *čergo-* 'to run' may be tentatively compared with Nostratic **t'äŕgE* 'to run' (*ŕ* = palatalized *r*) > Afroasiatic **√t'-r-g* > Berber **√ḍ-r-g* (*ḍ* = pharyngealized *d*) (Tayert, Tawellemet *ḍărrăg* 'to flee', etc.), Semitic **√t'-r-k'-* 'way, road' (Arabic, Tigre), possibly Chadic (Kera *təra* 'to run'); Indo-European **tragh-* or **trağh-* 'to run, move' (> Gothic *þragjan* 'to run', etc.); Altaic **t'äŕge* (*t'* = fortis *t*) > Turkic ** t'äŕ-* 'to run away, flee' (Old Turkish, Azeri *täz-*, Chuvash *tar-*, etc.), Mongolian *tergi-le-* id. A connection with Kurku *saṛub-* (proposed by Mundlay) is hardly acceptable. Kuiper (1966:76): a word of unidentified origin.

20. Nihali *jaluu-*

Nihali (Kuiper) *jaluu-* 'to descend', (Mundlay) *jaalú-* 'to climb down (from a tree)' (*j* = voiced palatal stop) may be very (!!) tentatively compared with Nostratic **zUlV* 'to fall down' > Kartvelian **źwel-/*źul-* 'to fall down' (Old Georgian *zul-eva-y* 'falling down' [verbal noun], Laz *žol* ~ *žul* 'to fall down'), Semitic **√z-l-l* 'to fall' (Jibbali *√z-l-l* id., Arabic *√z-l-l* 'to slip away'), (?) Indo-European **Hwel-/*Hwol-* 'to fall down, fell down' (Slavic **vali-ti* 'to fell down', Armenian *glem* 'to throw down'). Kuiper (1966:77): a word of unidentified origin.

21. Nihali _maa-_

Nihali (Mundlay) _maa-_, (Kuiper) _ma-_ ‘to give’ (according to Kuiper, a word of unidentified origin) may be tentatively compared with Nostratic _*mogV_ ‘to give as present’ > Indo-European _*megh-_ ‘to give’ (> Avestan _maga-_ ‘gift, grace’, Old Indian _mámhatē_ ‘he gives grants’), Semitic _*√m-g-n_ ‘to give as present, deliver, offer’ (> Ugaritic _√m-g-n_ ‘give presents’, Akkadian _magannu_, Aramaic _maggaan_ ‘gift, grace’, Phoenician _√m-g-n_ ‘to deliver, offer’, etc.), Dravidian _*moy-_ ‘gift, present’ (Tamil _moy_, etc. [cf. Burrow-Emeneau 1984:#5121]). An alternative possibility is that the Nihali word is an ancient loan from Dravidian.

22. Nihali _meŋge_

Nihali (Kuiper) _meŋge_ ‘tooth, jaw’, (Mundlay) _ménge_ ‘teeth’ resembles Nostratic _*mänt’V_ (or _*mant’V_) ‘jaw, chin’ > Indo-European _*m̥nto-_ ‘chin’ (Latin _mentum_, Welsh _mant_, etc.), Uralic (Samoyed: Nganasan _munduyśaŋ_, Nenets _munoćʔ_ ‘beard’), Dravidian _*maṇṭ-_ ‘skull’ (Tamil _maṇṭay_, etc.).

23. Nihali _oṭṭi_

Nihali (Kuiper) _oṭṭi_ ‘to burn (tr.)’ — Nostratic _*qot’U_ ‘to burn (tr., intr.), to kindle; fire’ > Semitic _*√x-t-w_ ‘to be kindled, burn’ (in Hebrew, Aramaic, Ge‘ez), Indo-European _*H_^{w}_et-_, _*H_^{w}_et-r-_ ‘fire’ (Avestan _ātarš_ id., Armenian _votër_ ‘hearth’, Irish _áith_ ‘stove’), Altaic _*oot‘a_ > Turkic _*ōt_ ‘fire’, Tungus (Evenki _otū_ ‘hearth’, etc.). According to Kuiper (1966:78), the Nihali word is of unidentified origin.

24. Nihali _páda_

Nihali (Mundlay) _páda_, (Kuiper) _paḍa_ ‘to kill’ may be tentatively compared with Nostratic _*p’at’a_ ‘to beat’ > Altaic _*p‘at‘V_ > Tungus _*pāt-_ ‘to strike, hit’, Turkic _*at-_ ‘to beat’; Semitic _*√p-t-t_ ‘to break into small pieces’ (in Hebrew, Syriac, Arabic, Ge‘ez); Kartvelian _*petk-_ ‘to explode, to beat (of heart)’.

25. Nihali _uri-_

Nihali (Kuiper) _uri-_ ‘to kindle’ (unless it is derived from Nihali [Mundlay] _ú-_ ‘to kindle’) may be compared with Nostratic _*warV_ (= _*wari?_)’ to kindle, to burn, to roast/fry on fire’ > Semitic _*√w-r-y_ (> Arabic _√w-r-y_ ‘to yield fire [steel], to be kindled [fire]’), Egyptian _wr.t_ ‘flame’, Chadic (Dera _wari_ ‘to roast’), Kartvelian _*warwar-_ (> Georgian _varvar-_ ‘to glow, to flame’, etc.), Indo-European _*wer-_ ‘to burn, kindle’ (> Armenian _varrem_ ‘I kindle’, _varrim_ ‘I burn’, etc.), Dravidian _*var̠-_ ‘to fry’ (> Tamil _var̠u_ ‘to be fried’, etc. [see Burrow-Emeneau 1984:#5325; Blažek 1992:#28]).

Conclusions

Of course, on the basis of this comparison, we cannot jump to any conclusions. Before drawing conclusions about the classification of Nihali, we must study the whole extant Nihali data (including Kuiper 1962, still unavailable to me), and check them on the basis of Indo-Aryan comparative phonology and etymology (Turner's dictionary) in order to exclude Aryan loans and to understand the phonologic prehistory of the languages of the area. Then, if the Nostratic (or Austroasiatic, or Sino-Caucasian) parallels are elucidated, we must establish regular sound correspondences and the phonetic laws standing behind them. Otherwise, any conclusions will be superficial and hardly convincing.

If the Nostratic origin of Nihali is the best hypothesis (as suggested so far by the pronouns), it shares with Dravidian a very important innovation in the realm of pronouns (see point 1), which suggests the existence of a Nihali-Dravidian (or Nihali-Elamo-Dravidian?) branch of Nostratic.

The Nihali-Munda lexical parallels, as presented by Mundlay, hardly prove the genetic relationship of Nihali with Munda. These lexical data fail to cover the central core of the vocabulary and the grammatical stock of phonemes. If these are the only shared words, they give the impression of going back to cultural contacts rather than to a genetic relationship.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Looking to the West and North: Nihali and Kusunda Find Links

Harold C. Fleming
Gloucester, Massachusetts, USA

Since international cooperative attempts at classifying Nihali (Nahali, Nehari), not to mention Kusunda, have barely begun, it is necessary to be quite explicit about basic data and methods. Also, I think that Nihali and Kusunda are positively related to each other and to most of the languages indigenous to western Eurasia and northeast Africa. It will require much hard work to develop fully the evidence for this conclusion, however. Let this summary be a contribution to the beginning of that evidentiary presentation.

At the very outset we must discuss the basic data. It is not the case that we are working with two languages whose dialects are lined up in rows, whose dictionaries with thousands of lexical entries are studded with etymologies, whose grammars are full and adequately exemplified, and whose exchanges with outside languages have been carefully evaluated. None of these desiderata are present. Not only that but an air of haste, yielding sloppiness, prevails in both corpora as presented by a series of authors. In the case of Kusunda, the world is faced with morbidity — the few remaining informants may well have died by now. In the case of Nihali, we face secrecy and the possibility that we are dealing with a "bandits' jargon". We feel lucky in both cases to have any firm data at all. Yet at the level of inner core vocabulary (basic vocabulary), both languages exhibit enough that appears to be native — enough for us to venture to classify, to relate to known outside groups and/or phyla.

A little simple counting is in order. The entire corpus of Kusunda which I know of and worked with is drawn from the work of Grierson and Reinhard. Even including the inevitable redundancy, the total corpus does not exceed 550 data entries (grammemes, words, phrases, sentences). For Nihali, we have the much richer corpus amassed by Mundlay, and presented in this volume, but also a substantial corpus from Bhattacharya and a lesser one from Grierson. All these Nihali sources, including much redundancy, do not exceed 2120 data entries. Moreover, while Reinhard has given a small amount of valuable grammar, our Nihali sources give next to nothing in this important domain. Kusunda verbal data are complex and poorly analyzed thus far.

Again one is reminded that the high standards of contemporary 'splitter' linguists have not been reached. Witness Terence Kaufman's "2000 good lexical items plus a good grammar before you can classify" (personal communication, 1990). Yet veteran taxonomists, as opposed to theoretical ones, will assure you that most languages in their areas have been successfully classified with *less data* than Kusunda offers. Where Kaufman's unreasonably high standards *do* become relevant is in the case of *remote or long range* genetic connections involving small phyla (isolates).

The problem of basic data is more than a question of poverty. We must ponder the inactivity of linguists and anthropologists who have known for *decades* that two highly distinctive and very endangered languages (and cultures) were expiring at their very feet, so to speak. What kinds of *priorities* were in effect that tempted linguists away from further research on them? Even research that they themselves had begun? Whatever those priorities may have been, it appears that capturing an important part of ancient Indian prehistory before it slipped away was *not* high on anyone's list. The often trivial and silly priorities of the 20th century underlie the data problem.

We need to mention *methods not used* first. Neither the famed mass comparison of Greenberg, nor the Muscovite reconstructionist approach, can be said to have been used. The demands of both are too much for preliminary hypothesizing to cope with. More explicitly, I and most of the other referees have not taken the time to line up large amounts of data systematically in order to do justice to those methods. Or we simply lacked the time.

Therefore, it is most important to deny that either Nihali or Kusunda has been rigorously evaluated in genetic terms. If it be said that no relations with outside groups were found, it must be added that the efforts fell short of the goals. To do mass comparison properly would demand comparing hundreds of languages. One colleague started to do just that but was unable to sustain the effort.

What we have relied on primarily is *word recognition* by those expert enough in some outside phyla. This approach means that in looking over data in a target language we notice some words — let me use *hypothetical* Indo-European examples — for example, **nasi* meaning 'nose' and **mi* for 'I' or 'we'. We recognize those as the same as, or akin to, typical Indo-European forms. We score these as potential links to Indo-European. We do not scorn them simply for being "look-alikes" or similarities, partly because they are not singular (links to just one language or one small group of a phylum) but plural (recurrent) and because it is silly to throw out potentially good evidence. However, when Basque, Sumerian, Burushaski, or Kusunda are involved, we are dealing with singularity, since each of these small phyla consists of one language or a cluster of dialects. There is nought we can do about that; singularity is less convincing than plurality. It always has been.

The pair mean different things also; **nasi* is virtually confined to Indo-European proper, while **mi* is found all over northern Eurasia and testifies to Nostratic links. If we were unable to find the first, we would still be encouraged by the second. Links to Nostratic are important in showing a more general or deeper relationship, while limited or exclusive links like **nasi* would have their value in arguing for specific classification (close to Indo-European) or in arguing that the 'nose' form is not limited to Indo-European.

Our hypothetical pair above illustrate the continuing taxonomic problem. Different pieces of evidence argue for different things. Many very close similarities exist between Nihali and the several phyla of linguistic India. The main mass of those look-alikes have been assigned to linguistic exchange — borrowing from someone. While past research has certainly cleared a lot of ground for us by pointing out the borrowings, there has been a tendency, perhaps excessive, to assume that Nihali borrowed while Indic or Dravidian lent. But in the case of Munda of Austroasiatic, there was a greater tendency to consider cognation as the explanation of similarities. A general bias or predisposition to perceive Nihali as somehow "primitive", "eastern", or the like — qualities assigned to Munda — has probably been involved. In Kusunda's case, its location in lower Himalaya predisposed one to see it as Sino-Tibetan or part of a hypothetical Munda stratum. Kusunda surely is neither of those.

Other pieces of evidence argue for various things. A language being newly inducted into a phylum — like Ongota recently in Afrasian — will show cognates with various parts of the phylum, almost at random. If it nevertheless shows preferences (more cognates) for a particular group, then its sub-classification within the phylum is indicated. If the language remains even-handed as it were, sharing cognates about equally with various sub-groups, then its own status as a sub-group becomes apparent. It relates to the whole but not to any particular part, so it itself is a branch or distinct sub-group.

But the pieces have greater sublety than that. Suppose that the various groups being compared have great differences in membership. One group is internally complex with many members; the other has few. Then one has to expect the complex group to manifest more cognates than the homogeneous one because complexity preserves a much greater amount of prehistoric data. Consequently, one has to expect, all other things being equal, that a homogeneous group will be harder to classify because it preserves *less evidence* of relationship. Not only is that the continuing problem with Basque, Sumerian, Burushic, etc., it still is for Albanian as opposed to Romance or Indic in Indo-European or Brahui compared to peninsular Dravidian. Were Egyptian not backed up by 3000+ years of written records, it would be difficult to assess, as opposed to the very complex Chadic group of Afrasian.

The final aspect of methods is what was compared. Relying on colleagues to cover areas of their own expertise, I did not cover the vast realms of Nostratic or Dene-Caucasic. Some of their cognates were found but only because of their familiarity. Semi-systematic searches of basic vocabulary in Basque and Burushic were made but were limited by my own inadequate compilation of comparable Basque or Burushic data. My most important contribution was the Afrasian viewpoint, where my own expertise lies, especially in Cushitic, Omotic, and Ongotan. No use was made of published reconstructions of Afrasian or Nilo-Saharan. Neither Niger-Congo nor Khoisan were considered in a serious way. None of the linguistic phyla to the *east* of central India were compared; again I relied on knowledgeable colleagues for that. My impression is that the Austric realm will eventually produce links to Nihali.

What are the Results?

As the roughly 30 cognations presented below will argue, Nihali finds some cognations, all convincing I hope, with phyla to the west and northwest of central India. Moreover, it has a few strong links to Kusunda to its north, while Kusunda has its own links to the west. Nihali has at least one link to two inner African phyla — Koman cum Shabo of Nilo-Saharan and Khoisan — which lose their exciting latent exclusivity because they are part of a *global* or nearly global cognate. This is not to deny that the inner African phyla might turn up other cognates; I did not examine them systematically.

Pronouns:

1. 'I' 1st person singular:

Nihali *joo / juo / jo*
Kusunda *či / tsi / čiyi*. Also *ton-da* 'my'. One occurence of *ki*, which is not the source of the other forms.
Burushic *ja*
Ongota (Afrasian) ('we') *jo / jo-ta*
West Caucasic *sa- / sê*
East Caucasic *su / sun / zu / zun*, etc.

(Ruhlen, 1995:254-56 has these additional forms, below)
Austroasiatic *joo* (1 of 3 forms)
Hattic *se / es* (See W.Caucasic)
Hurrian-Urartean *se- / es- / yešâ / šo* (1st of two forms)
Yeniseian **ʔaj*
Na-Dene: proto-Athabascan **šwíí*
Khoisan 'we (exclusive)' **si / sii* (there are 2 inclusive forms)

2. 'we' 1st person plural:

Nihali *téko / ṭeeku* 'we two'
Kusunda *toʔi / čo / tok*
West Caucasic *t- / d-* (Circassian only)
East Caucasic *čun / čin / uču* (Samurian group only)
Kartvelian *čkhi / čkhin / čven / čon* (but not Svan)
North Omotic (Afrasian) 'I' *ta / ta-na* (but not Dizoid)
Chadic (Afrasian) *tem / dem / dann* (scattered and isolated)

(Ruhlen, 1995:254-6 has these additional forms, below)
Khoisan 'I' *tii* (1 of 3 forms)
Indo-Pacific *ti* (1 of 4 forms)

3. 'thou' Second person singular:

Nihali *nee / nye*
Kusunda *nu / nû / ni* (thy) / *niyi* (thy)
Burushic (?) *un / uŋ*
North Omotic *ne / ne-na*

(Ruhlen, 1995:257-9 has these additional forms, below)
Dravidian *nii-ṇ / nyi-ṇ* (1 of 2 forms; proposed loan to Nihali)
Elamitic *ni / nin* (1 of 2 forms)
Sino-Tibetan **na / naŋ /nyo*
Na-Dene: proto-Athabascan *nani / nine / niŋ / nì*
Korean-Japanese-Ainu *na / nâ* (1 of 2 forms)
Nilo-Saharan **ini*
Indo-Pacific *na / ni / ŋi* (1 of 4 forms)

Australian **ŋin / ŋiñ*
Kordofanian **ŋa / ŋô / ŋo*

4. 'you, ye' Second person plural. (Lines 3 and 4 from Ruhlen as cited):

Nihali *nyeko / naaku* (you two)
Kusunda *noʔi / noki*
Austroasiatic **inaa* (1 of 4 forms)
Indo-Pacific *nik* (1 of 4)
Miao-Yao *ńew* (1 of 2 forms. Not a plural in form, so cf 'thou')

Inner Core Vocabulary

5. 'blood, flow of blood':

Nihali *apse / apse-bando* 'to check flow of blood' (As cited)
Afrasian: Chadic: Proto-Chadic (Jungraithmayr) **bz* / *b_2z* (Found mostly in Central Chadic but also in Eastern)

6. 'blood vessel, veins':

Nihali *siraa*
Afrasian 'root, vein, nerve': Greenberg #60, e.g., Semitic Amharic *sîr*, Hebrew *šêrêš*, Berber Tuareg *azar*, *asur*, etc.

7. 'to blossom, to flower, ; flower':

Nihali *phúl*
Proto-Chadic (Jungraithmayr) **pl*

8. 'breast (woman's)':

Kusunda *ambu*
Afrasian 'breast (&): Ongota *ʔaama*, Cushitic: Tsamai *ʕamb'e*, Gawwada *ʕamo*, Konso *ama*, Gidole, Bussa *amp'a*, Somotic: Ari *aami*, Dime *ʔîmmo*, Nomotic: Mao *ammi* (in other branches too)

9. 'mother's brother, father's younger sister's husband / MoBrWi':

Nihali *maama / maami*
Kusunda *mâm* ElBr, older brother'
Afrasian: Cushitic Dahalo 'MoBr' *ʔáám̯a*, Semitic Arabic 'FaBr' *ʕam*

10. 'brother, younger':

Nihali *sanu*
Afrasian 'brother(1)' Greenberg #13, e.g., Egyptian *sn*, Beja *san*, Somrai *sen*
(Also found outside of Afrasian)

11. 'to burn (intransitive)':

Nihali *aḍé- / aḍek-*
Burushic *âtsi-âs*
Afrasian: common Omotic, e.g., Kafa at'-, Ometo *âts' / eč' / et*; Dizi *at- / ats-*, Somotic Ari *ats'*, Hamar at, Dime *ats'*

12. ‘to cry’:

Nihali *apa*
Afrasian: common Nomotic *ef*

13. ‘to chew, to eat’:

Nihali *té / ṭee-*
If this is one verb, heard differently, then the 1st form has Afrasian and Nostratic cognates. If the retroflex 2nd form is separate, then it has Afrasian cognates. For the 1st form see Greenberg #27 ‘to eat’. E.g., Berber *ca / tett*, Egyptian ‘bread’ *t3*, Semitic Akkadian *teʔu*. For the 2nd form, see Semitic Arabic ‘bite’ *ʕaḍḍ*, Omotic ‘tooth’ **as'*, Cushitic Burunge ‘tooth’ *atl'-imo / atl'-o* (pl).

14. ‘eye / to see’:

Kusunda *tsan-*
Burushic *-lcin* / — (Problem [1])

15. ‘eye(ball)’:

Kusunda ‘eye’ *inin / iniŋ* (usually cited with prefix [tai-])
Burushic *nâna / -lcin-ê nâna* (as cited = eye-of ball)

16. ‘fire’:

Nihali *aapo*
Burushic *pfu*

17. ‘to fly / fire-fly’:

Nihali *aphir / aphiri*. The striking correlation between these two seems natural; it also occurs elsewhere, especially in Omotic. 1st form may be borrowed from Korku but as part of a global etymology (Ruhlen, 1995. #19) it need not be due to borrowing.

18. ‘foot’:

Kusunda *iyan / wân / tai-wân-gepan* ‘toe’ (as cited = ‘tai-foot-digit’)
Basque *oyn*

19 ‘to give’:

Nihali *bee-*
Afrasian: Semitic Arabic *byʕ*; Ongota *biʔe*
Nilo-Saharan: Surma Kwegu (rare form) *búuà* ‘hand’; Nyangeya (Kuliak) *bêê* ‘give’; Tepeth (Kuliak) *miyaʔ* ‘to give’; Ik (Kuliak) *mê-ês* ‘give’

20. ‘to give’:

Nihali *ma- / maa-*
Afrasian: common Omotic *im-*; Ancient Egyptian *ỉmỉ* (not regular verb).
See Kuliak forms above for probable cognates.

21. ‘good, beautiful’:

Kusunda *weyn / wâyyâki / wei - wâyâi*
Basque *ona* ‘good’.
Proto-Nostratic (Bomhard) **win-/*wen-* ‘to strive for, wish for, desire’. This is the root of modern English ‘win’ rather than ‘want’ which is from PIE **wā-no-*. Also Ongota *wanna* ‘good’, Somotic Ari *waanna* ‘good’; Cushitic: Dahaloan ‘good, beautiful, clean’ *wíne / více / vino*; Iraqw ‘soft, gentle’ *wanana*; proto-South Cushitic (Ehret, 1980) **win-/*wan-* ‘nice, pleasant, comfortable’. Bomhard (1994) relates these forms in turn to Egyptian *wnf* ‘be joyful, rejoice’; and to his proto-Afroasiatic **wan-* ‘be joyful, pleasant’. Nihali has *awol* ‘good’ and *awalka* ‘good, better, best’; they might be cognate to Kusunda *weyn*.

22. ‘louse’:

Nihali *se / si*
Afrasian: Beja Bisharin ‘camel louse’ *se*, Hadendiwa ‘tick’ *sâʔi-t*, Cushitic Ma’a ‘tick on an animal’ *swaʔ*, Tsamai ‘flea’ *sae*; Somotic Ari ‘flea’ *saya / sea* and ‘tick’ *šoʔ*; Galila ‘flea’ *seya* and ‘tick’ *šay*; Dime ‘flea’ *sîye*; Nomotic Dizi ‘flea’ *tsáhi*.
Nilo-Saharan: Nile Nubian / Dair *issi / itu*; Nilotic Masai *ašei*.

23. ‘male’:

Nihali *jaakoṭo*
Afrasian: Ongota *šoqta* ‘male (animal)’ Singularity weakens the case, even if Ongotan is a branch.

24. ‘mother’:

Nihali *aay* — possibly borrowed from Marathi or vice versa; because this is a global etymology. While nearly as common as ‘mama’ or ‘dada’, it cannot be derived from baby talk.

25. ‘nose, hole, mouth’ (Dialect differences. Primary meaning = ‘orifice’ or the like):

Kusunda *aau / au*
Basque *aho / au* ‘mouth’

26. ‘nose’:

Nihali *coon / cón / cojon / cón-dhanḍi / cón-popa*. Three field workers support the 1st form. One lists the 3rd form as distinct. Some confusion caused by occasional appearance of *coʔon* in the literature but in nobody’s field notes. The 4th and 5th forms mean ‘bridge of nose’ and ‘nostril’ respectively. They also support *coon*. This is the form listed in Ruhlen, 1995, #4, as **čuna / čunga* ‘nose, to smell’. While the overall presentation has great force in its representation of 18 phyla, errors occur with some frequency too. The remarkable similarity of recurring versions of the old base in Omotic and Amerind suggest strongly that the initial consonant was either **s-* or the retroflex **ś-*. This is supported by both Shabo of Africa and Nihali of India via recurrent examples of [s] and [č] alternating initially. Moreover, the recurrent world-wide forms strongly suggest a tripartite morpheme alternatively **sina / *sint’a / *sink’a* > *singa* and *siŋ*.

Quite unprincipled attempts by some scholars to clip the final consonants off two of these forms by treating them as bound forms (for which there is no evidence) should be rejected. Trask's criticisms of opportunistic segmentation are remembered. Let us not reduce *sint'* of Omotic to *sin-t* just because we want the base to be *sin* only. In card playing that is called 'cheating'. In archeology, one might carve artifacts to make them more agreeable! Nihali has a verb or verbal noun *sani* with meaning 'smell' assigned to it; it was buried in the data on an insult word [caapiniku]. I presume that *sani* comes from the same base as *coon*, so it is not compared by itself. By itself *sani* could involve Nihali in the global etymology. Also noteworthy is that the Kusunda data do not include any forms related to *coon*. Kusunda 'nose' is *(tai)-nao*.

27. 'one (cardinal number)':

Nihali *bad̤aa / bed̤e / bid̤i*
Basque *bat*

28. 'rat':

Nihali *hond̤ar*
Afrasian: Nomotic forms in *iins'a / iind'i / iinco* but local borrowing may explain it.

29. 'to see, to examine closely':

Nihali *araay / ará*
Afrasian: *ar-*, *arx-* 'to see, know' pervades Afrasian.

30. 'uncle, aunt: MoYoSiHu, FaSiHu, FaYoBr, etc./ FaYoBrWi':

Nihali *kaka / kaki*
Afrasian: 'grandparent, ancestor'; Cushitic: Proto-Highland East Cushitic **akaako* (Hudson), Gollango *akká-wo / akká-ye* Somotic: Kara 'GrMo' *kaaká*, Hamar 'GrFa / GrMo' *êke / aka*; Nomotic: Dizi (Adikas) 'FaFaFa' *aku*; Ongota 'GrFa = FaElBr, FaFa, MoFa' *ʔakka* (it may be borrowed from Cushitic Tsamai *ʔakka*) (See Ruhlen, 1995:218-9 for etymology #1 'older brother', where the best match for Nihali is in Amerind!)

31 'who?, what?, why?':

Nihali *naani / naanko / naavay*. The interrogative base seems to be *naa-*.
Kusunda *nâtâi / nâtât / nâtan / nâtaŋ / nâtia*. The base appears to be *nât-* 'who? / who? / what? / what? / why?' Compare Ongota *neení /niike* 'what?'; Beja Hadareb (Hadendiwa) *nân* 'what?', Bisharin *na / nan* 'what?'. Nilo-Saharan: Nubian *naa / nii/ naai / naandi* 'what, which, why?'; Kunama *naa*; Nara *na*; Nilotic Dinka *ŋa*, Shilluk *aŋo*; Saharan Teda *aŋa*: all mean 'what?' Shabo has *ne / ne-ge* 'who? / who?-plural.

Grammatical Evidence

Nihali's morphology, being underdescribed and difficult to determine, is left for the future. For now we can say that two obvious grammemes of widespread occurence are also found in Nihali, to wit:

Genitive (possessive) in *kaa / kee / ku*, usually noun suffixes, and genitive in *-n / -ne*, also usually noun suffixes. The [k] is found in Afrasian at least, while the [n] is found there and in Basque and many other languages. At present, the genitive in Kusunda cannot be determined.

However, some strong verb conjugation data exist which show that Kusunda has *t-* / *n-* / *g-* for 1st / 2nd / 3rd persons as verb prefixes. Not only do these confirm the focal pronouns, but the pattern as a whole is remarkably similar to Nomotic generally, but also especially to Somotic Dime's suffixed pattern.

General Conclusions

No doubt the Indian Ocean is somewhat challenging to sailors. Yet their navigations on it have continually brought trade goods and peoples from southern Asia and eastern Africa into contact. Still our linguistic community has been a prisoner to geography. No doubt the efforts of the past to classify Nihali and Kusunda have only walked to the end of the block, so to speak, but not looked across the sea. Else why were the clear and strong links to Africa not noticed before? Indeed, one may bet that the African links are stronger than those to southeast Asia and Oceania, despite the fixation on Austric and notions of "Australoid aborigines" being the basal autochthones of greater India. One must remember that India was first settled from the West, if the African homeland hypothesis is correct, and that the earliest autochthones had come from Africa, not Australia. The language taxonomy of India's moribund and disparate groups seems to bear out that point. And yet, given India's role as the cultural diffusion cul-de-sac of Asia, alongside its mission as radiator of religion, we may have missed the lower strata of prehistory after all!

ADDENA: Thanks to John Bengtson: two new sources on Kusunda. One, by N. V. Gurov of St. Petersburg, Russia: "Kusunda sinokavkazskie leksičeskie paralleli (k xarak-terisike načal'nogo etapa formirovanija Južnoaziats-kogo jazykovogo areala)" in *Lingvističeskaja rekonstrukcija i drevnejškaja istorija vostoka.* čast 3. Moskva: Institut Vostokovedenija 1989, pp. 41-44. No translation into English has appeared yet. Gurov lists 20 etymologies which seem to link Kusunda with Caucasic and Yeniseian, perhaps others. Some are convincing. However, he did not use Reinhard's data, unfortunately. Two, Reinhard (*JINAS* 4/1:1-21 [1976]) has a Swadesh list of Kusunda; I have not seen that. Some of the data conflict with his publication with Toba, especially whether [taa] or [taai] is a prefix and just what 'eye' is — *činiŋ* or *č-iniŋ* or *inin*!

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Nihali and Austroasiatic

Ilia Peiros
Department of Linguistics and Applied Linguistics
The University of Melbourne
Parkville, VIC 3052, Australia

Two languages, A and B, are genetically related if the following conditions are fulfilled:

Condition I: The Existence of Similar Morphemes:

(i) Genetically related languages A and B share a sufficient number of similar morphemes.

If the two languages reveal a sufficient number of similar morphemes, one can assume that this similarity is not accidental, and it can probably be interpreted as evidence of genetic relationship. Similar morphemes can be either lexical or grammatical, but the existence of similar lexical morphemes seems to be obligatory: there is no generally accepted language family for which languages do not share similar lexical morphemes, while for several well-established families of Southeast Asia (Kadai, Vietic, Lolo-Burmese), no grammatical morphemes are known. If similar morphemes are not found, we do not have data for further discussion of genetic relationship of these languages.

Condition II: Genetic Reasons for the Similarities between Morphemes:

(ii) A sufficient number of similar morphemes in languages A and B belongs to the core lexicon.

Similar morphemes can be found in all parts of the lexicon, but if the languages are really related, they always share morphemes from the "core lexicon". We will develop and discuss this notion later. Here, it is enough to say that core lexicon includes words with simple universal meanings, which are less open to borrowing than other parts of a language's lexicon. There is no doubt that words from the core lexicon can also be borrowed, but the likelihood of borrowing here is usually lower. To the best of my knowledge, all known related languages always share words from the core lexicon. Thus, one can conclude, that if two languages share not only similar morphemes but these morphemes also belong to the core lexicon, then it is more probable that these languages are genetically related.

Condition III: The Existence of Systematic[1] Phonological Correspondences:

(iii) The phonological systems of A and B are connected by systematic phonological correspondences, an element of one system corresponding to certain elements (one, several, or none) in another.

(iv) The systematic phonological correspondences mentioned in (iii) are true for lexical similarities discussed in (i) and (ii).

Similarities between morphemes can be due to various causes: common origin, borrowing, chance resemblances, and so on. To demonstrate the genetic nature of these similarities, we need a system of phonological correspondences between the languages.

---

[1] A phonological correspondence is a systematic one if it brings together reflexes of a particular proto-phoneme. By definition, a correspondence based on borrowings cannot be systematic.

These conditions are sufficient to provide us with formal criteria to judge if there is enough evidence to accept that two languages are genetically related and (due to transitivity of the notion) that all languages related to them belong to the same linguistic family. It is important to mention that there is no additional requirement for grammatical similarities, as they are not universal, and languages in their development can lose most or even all of their original morphological properties. However, where there are such similarities, they can provide an additional and often crucial support for a genetic claim.

In many cases, we do not have phonological correspondences, and our judgment about the genetic affiliations of a language has to be based only on its lexicon.[2] Such judgments can be accepted, however, only as preliminary considerations.

It is well known also that comparative linguistics does not have procedures which allow us to demonstrate that two languages are not genetically related. Studying languages, we can only demonstrate that they are genetically related, or we can admit that there is not sufficient evidence to confirm their common origin.

From the above, it follows that the main precondition for any investigation into genetic relationship of languages is based on a list of similar morphemes (words) found in the languages under consideration.

A word list of Nihali by Asha Mundlay includes four different types of words similar to those identified by Kuiper (1962; 1966).

1. Words borrowed from Indo-Aryan languages;
2. Words borrowed from Dravidian languages;
3. Words also found in Munda languages;
4. Words with no external comparisons.

These four groups are also represented in the (unfortunately incomplete) lexicostatistical list of Nihali.

In the list, only nine Nihali words are similar to those of Munda:

- One Nihali word ('come') is found only in Sora;
- One Nihali word ('breast') is found only in Korku;
- One Nihali word ('knee') is found only in Mundari, but in both languages, it can be a borrowing from an Indo-Aryan source;
- Only one Nihali word ('tongue') is found in all four Munda languages included on the list;
- Three Nihali words ('dog', 'fly', and 'sand') are found in Mundari and Korku;
- Three Nihali words ('bark', 'give', and 'we') are found only in Nihali and Bonda.

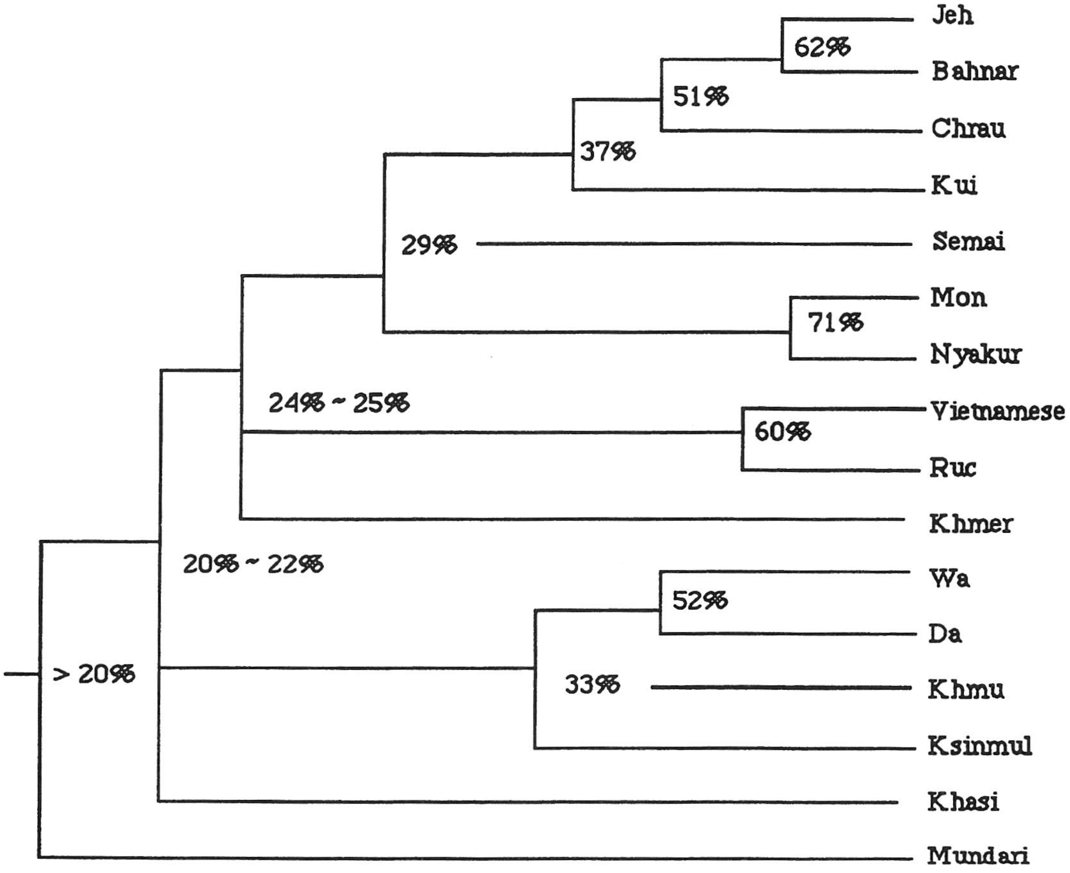
The number of the comparisons and the distribution of Munda forms make the hypothesis of a Nihali ~ Munda relationship not very convincing.

It is generally accepted that Munda languages belong to the Austroasiatic language family. Its internal structure is not yet absolutely clear, as a detailed Austroasiatic reconstruction is still missing. My provisional lexicostatistical classification of the family (Peiros, in press) is given below:

---

[2] Typological and other non-genetic considerations cannot be used here.

A Lexicostatistical Classification of Austroasiatic Languages



Unfortunately, this lexicostatistical classification is based at several points on differences of 3 or 4 percent, which makes it somewhat speculative. It is interesting that there is no clear distinction between Mon-Khmer languages (Central, Viet-Muong, Khmer, and Khasi) and Mundari of the Munda family, which indicates the possibility that there is no difference between the Mon-Khmer and Austroasiatic levels.

If Nihali is genetically related to Munda, it has to share some words with other Austroasiatic languages. But in the list given below, I have identified only two possible similarities of this type, 'sand' and 'what', a fact which demonstrates that we have no reasons to include Nihali in the Austroasiatic family and that its similarities with Munda are the results of borrowing.

As we also have no evidence to believe that Nihali is related to Dravidian, we have no other option than to assume that the language remains genetically isolated.

Appendix:
A Lexicostatistical List for Nihali

1. **all (we all ...)**

Nihali	==	0
Bonda	gulay	1
Sora	kúdub	2
Mundari	soben	L
Korku	par	3
Mon	klaʔ	4
Khmer	daʔ	5

2. **ashes**

Nihali	néto	1
Bonda	ug-sɔŋ	2
Sora	kúmab	3
Mundari	toroeʔ	4
Korku	hob	5
Mon	pateŋ	6
Khmer	phɛ:h	7

3. **bark (n.)**

Nihali	sáli	1
Bonda	úsa	1
Sora	kúrraŋ	2
Mundari	harta	3
Korku	==	0
Mon	sakow	2
Khmer	==	0

4. **belly**

Nihali	==	0
Bonda	sulɔi	1
Sora	kəmpúŋ	2
Mundari	laʔi	3
Korku	laʒ	3
Mon	buŋ	4
Khmer	baoh	5

5. **big**

Nihali	baṛa:	L
Bonda	munaʔ	1
Sora	sʊda:	2
Mundari	maraŋ	3
Korku	khad	4
Mon	==	0
Khmer	dho:m	5

6. **bird**

Nihali	poy	1
Bonda	piriʔ	2
Sora	ontíd	L
Mundari	ceṇe	L
Korku	==	0
Mon	gacem	3
Khmer	sat-sla:p	4

7. **bite**

Nihali	ha:ru	1
Bonda	ɔp	2
Sora	ram-	3
Mundari	hua	2
Korku	kab	2
Mon	kit	4
Khmer	kha:m	5

8. **black**

Nihali	==	0
Bonda	miṛeʔ	1
Sora	yága:	2
Mundari	hende	3
Korku	kende	3
Mon	caok	4
Khmer	khəmaw	5

9. blood

Nihali	corṭo	1
Bonda	boni	2
Sora	míɲam	3
Mundari	mayam	3
Korku	mayum	3
Mon	chim	3
Khmer	ʒha:m	3

10. bone

Nihali	==	0
Bonda	siʔ-saŋ	1
Sora	ə-ʒá:ŋ	1
Mundari	ʒaŋ	1
Korku	haṛge	L
Mon	cut	2
Khmer	chəʔɨŋ	1

11. breast

Nihali	buču	1
Bonda	daʔtukui	2
Sora	memé:	3
Mundari	nunu	4
Korku	bucu	1
Mon	tah	5
Khmer	taoh	5

12. burn (v.t.)

Nihali	law	L
Bonda	ɔgep	2
Sora	soi-	
	bal-	3
Mundari	ro	4
Korku	ǯul, ǯuluʔ	5
Mon	tau	6
Khmer	ṭut	7

13. claw

Nihali	==	0
Bonda	krimei	1
Sora	tar	2
Mundari	rama	3
	sarsar	
Korku	==	0
Mon	saṇem	4
Khmer	krəco:k	5

14. cloud

Nihali	==	0
Bonda	tirib=ḍak	1
Sora	tárub	1
Mundari	rimbil	2
Korku	==	0
Mon	mat.brai	4
Khmer	bəbo:k	5

15. cold (as water)

Nihali	==	0
Bonda	sep	1
Sora	ráŋa:	2
Mundari	tutukun	3
Korku	raban, rabaŋ	4
Mon	boeh	5
Khmer	treʒak	6

16. come

Nihali	páto	
	iyer	1
Bonda	sap-	2
Sora	ir-a:i	1
Mundai	hiʒuʔ	3
Korku	heǯ, haǯ	3
Mon	kluŋ	4
Khmer	mo:k	5

17. **die**

Nihali	beṭṭo	1
	ra:wanbi	
Bonda	gei-	2
Sora	kayéd	3
Mundari	goʒoʔ	2
Mon	khyuit	3
Khmer	slap	4

18. **dog**

Nihali	sita:	1
Bonda	gusɔʔ	2
Sora	kinsó:ʀ, kinsó:d	2
Mundai	seta	1
Korku	sita	1
Mon	kluiw	3
Khmer	chəkɛ:	4

19. **drink**

Nihali	==	0
Bonda	uʔ	1
Sora	ga-	2
Mundari	nu	3
Korku	nu	3
Mon	suŋ	5
Khmer	phɨk	6

20. **dry**

Nihali	patari	1
Bonda	nsɔr	2
Sora	asár	2
Mundari	roṛo	3
Korku	lok(h)or	4
Mon	parom	5
Khmer	sŋuət	6

21. **ear**

Nihali	ciga:m	1
Bonda	luntur	2
Sora	lʔu:d	2
Mundari	lutur	2
Korku	lutur	2
Mon	katow	2
Khmer	trəciək	3

22. **earth**

Nihali	kha:ra:	1
Bonda	tubuk	2
Sora	ləbó:	3
Mundari	ote	4
Korku	kasa	5
Mon	ti	4
Khmer	ṭi:	4

23. **eat**

Nihali	==	0
Bonda	sum-	1
Sora	ʒum	1
Mundari	ʒom	1
Korku	ʒom	1
Mon	ca	2
Khmer	ŋa:m	3

24. **egg**

Nihali	kalen	1
Bonda	ntɔʔ	2
Sora	ár-re	3
Mundari	ʒarom	4
Korku	atkom	5
Mon	khamhay	6
Khmer	bo:ŋ	7

25. **eye**

Nihali	ʒiki	1
Bonda	mɔʔ	2
Sora	mʔo:d	2
Mundari	med	2
Korku	med	2
Mon	mat	2
Khmer	bhənɛ:k	3

26. **fat (n.)**

Nihali	==	0
Bonda	kiri	1
Sora	kári:	1
Mundari	itil	2
Korku	carbi, cerbi	L
Mon	==	0
Khmer	==	0

27. **feather**

Nihali	poye	1
Bonda	bileʔ	2
Sora	balé:d	2
Mundari	iil	3
Korku	katha	4
Mon	==	0
Khmer	slap	5

28. **fire**

Nihali	==	0
Bonda	suŋɔ	1
Sora	tʔóge:	2
Mundari	seŋgel	2
Korku	siŋgel	2
Mon	pamat	3
Khmer	bhlaəŋ	4

29. **fish**

Nihali	cán	1
Bonda	aṛɔŋ	2
Sora	ə-yo:	3
Mundari	ha=i	4
Korku	ka=ku	4
Mon	ka	4
Khmer	tri:	5

30. **fly (v.)**

Nihali	aphir	1
Bonda	uṛ	L
Sora	e:ŋ-	2
Mundari	apir	1
Korku	aphir	1
Mon	==	0
Khmer	haər	3

31. **foot**

Nihali	==	0
Bonda	tiksuŋ	1
Sora	ʒʔe:ŋ	2
Mundari	kaṭa	3
Korku	naŋga	4
Mon	ʒoŋ	2
Khmer	ʒaəŋ	2

32. **full**

Nihali	bherya	L
Bonda	==	0
Sora	bariʒ	L
Mundari	pura	L
Korku	==	0
Mon	peŋ	1
Khmer	bə:ŋ	1

33. **give**

Nihali	bé	1
Bonda	beʔ	1
Sora	tiy-	2
Mundari	om	3
Korku	i, o	4
	ʒii, ʒuuʔ	
Mon	kuiw	5
Khmer	aoj	6

34. **good**

Nihali	awol	1
Bonda	bani	2
Sora	báŋsa:	3
Mundari	bugi	4
Korku	(b)ain, ayn	L
Mon	khuih	5
Khmer	ʔəʔo:	6

35. green

Nihali	==	0
Bonda	==	0
Sora	(óla:-)gən	1
Mundari	gaded	2
Korku	nila, lila	L
Mon	saŋeak	3
Khmer	khiəw	L2

36. hair (of head)

Nihali	kugusu	1
Bonda	ug=bɔk	2
Sora	kúru	3
Mundari	ub	4
Korku	hub	4
Mon	sok	2
Khmer	sok	2

37. hand

Nihali	hat	L
Bonda	titi	1
Sora	sʔi:	1
Mundari	ti	1
Korku	tii	1
Mon	tay	1
Khmer	ṭaj	1

38. head

Nihali	pyeŋ	1
Bonda	==	0
Sora	bʔo:b	2
Mundari	booʔ	2
Korku	kapar	L
Mon	kaḍuip	3
Khmer	kəpa:l	L

39. hear

Nihali	ca:kini, cekini	1
Bonda	ɔŋ-	2
Sora	amdáŋ	3
Mundari	ayum	4
Korku	aɲʒum	4
Mon	miŋ	5
Khmer	lɨ:	6

40. heart

Nihali	==	0
Bonda	buruk	1
Sora	púra:da:n	L
Mundari	ʒi kuṛam	L2+3
Korku	==	0
Mon	gruih	2
Khmer	pəhṭu:ŋ	L3

41. horn

Nihali	singi	L
Bonda	ḍeruŋ	1
Sora	dére:ŋ	1
Mundari	diriŋ	1
Korku	siŋgi	L
Mon	draŋ	1
Khmer	sənɛ:ŋ	2

42. I

Nihali	jó	1
Bonda	niŋ	2
Sora	ɲen, nyen	2
Mundari	aiɲ	3
Korku	aɲ	3
Mon	ʔay	3
Khmer	khəɲom	4

43. kill

Nihali	páda	1
Bonda	ɔgɔi	2
Sora	káb-yed-	3
Mundari	goeʔ	2
Korku	goʒ	2
Mon	gacuit	2
Khmer	səmlap	4

44. knee

Nihali	mokh(a)ne	1
Bonda	manḍi	L
Sora	mándidi	L
Mundari	mukuṛi	1
Korku	topare	2
Mon	kboaŋ	3
Khmer	ʒəŋgaŋ	4

45. know

Nihali	==	0
Bonda	mak-	1
Sora	gálam-	2
Mundari	itu	3
Korku	cina malum	4
Mon	tim	5
Khmer	ṭɨŋ	6

46. leaf

Nihali	cokob	1
Bonda	paʔṛu ulak	2
Sora	ó:la:	2
Mundari	sakam	3
Korku	==	0
Mon	sla	2
Khmer	səlɨk	4

47. lie (down)

Nihali	==	0
Bonda	==	0
Sora	lud-	1
Mundari	gitiʔ	2
Korku	luriʒ	1
Mon	wat	3
Khmer	dəmrə:t	4

48. liver

Nihali	==	0
Bonda	gire	1
Sora	gáre:	1
Mundari	iim	2
Korku	koleɲʒa	L
Mon	kta:n	3
Khmer	thlaəm	4

49. long (as rope)

Nihali	bha:geli	1
Bonda	sileŋ	2
Sora	ʒalé:	2
Mundari	ʒiliŋ	2
Korku	ghiliŋ	2
Mon	ʒliŋ	2
Khmer	vɛ:ŋ	3

50. louse (hair)

Nihali	cilar-ṭa kepa:	1
Bonda	gisi	2
Sora	iʔi:	2
Mundari	si-ku	2
Korku	si-ku	2
Mon	cay	2
Khmer	caj	2

51. **man**

Nihali	==	0
Bonda	remɔ	1
Sora	mán(d)ra:	L
Mundari	koṛa	2
Korku	koro	2
Mon	truʔ	3
Khmer	proh	3

52. **many**

Nihali	khobo	L
Bonda	gulay	1
Sora	boi-bói	2
Mundari	bese	L
Korku	ghoneǯ leka	3
Mon	gluiŋ	4
Khmer	craən	5

53. **meat**

Nihali	ka:w	1
Bonda	sili, seli	2
Sora	ʒélu:	2
Mundari	ʒilu	2
Korku	ǯilu'	2
Mon	cun	3
Khmer	sac	4

54. **moon**

Nihali	cended, thénded	L
Bonda	arke	2
Sora	aŋgáʒ	3
Mundari	caṇḍuʔ	L
Korku	==	0
Mon	gatu	4
Khmer	khɛ:	5

55. **mountain**

Nihali	balla:	L
Bonda	kunḍa	L2
Sora	barú:	1
Mundari	buru	1
Korku	balla	L
Mon	duiw	2
Khmer	bhəno:m	3

56. **month**

Nihali	ka:ggo	1
Bonda	tumɔʔ	2
Sora	tʔo:d	2
Mundari	moca	3
Korku	cabu	4
Mon	paŋ	5
Khmer	mat	2

57. **name**

Nihali	==	0
Bonda	nimi	1
Sora	əɲóm	1
Mundari	nutum	1
Korku	ʒumu, ʒimu	2
Mon	imu, yamu	1
Khmer	ʒhəmoah	2

58. **neck**

Nihali	gardan	L
Bonda	gulubup	1
Sora	sáŋka:	2
Mundari	hoṭoʔ	3
Korku	==	0
Mon	kaʔ	4
Khmer	kɔ:	4

59. **new**

Nihali	nava:	L
Bonda	time	1
Sora	tamme	1
Mundari	nawa	L
Korku	==	0
Mon	tami	1
Khmer	thəmi:	1

60. **night**

Nihali	méndi	L
Bonda	tumuŋgɔ	1
Sora	tó-gal	2
Mundari	nida	L2
Korku	rato	L3
Mon	bətami	3
Khmer	jop	4

61. **no, not**

Nihali	==	0
Bonda	anḍa	1
Sora	íʒʒa	2
Mundari	ka	3
Korku	==	0
Mon	ha, hwa	4
Khmer	ʔɔt	5

62. **nose**

Nihali	cojona	1
Bonda	nseʔmiʔ	2
Sora	mʔu:	2
Mundari	mu~	2
Korku	muu	2
Mon	muh	2
Khmer	crəmuh	2

63. **one**

Nihali	baḍa:, biḍiko	1
Bonda	muiḍa	2
Sora	əbói	3
Mundari	miyod, mid	2
Korku	miyaʔ	2
Mon	moy	2
Khmer	muəj	2

64. **person**

Nihali	jana:	L
Bonda	==	0
Sora	mándra:	L2
Mundari	hoṛo	1
Korku	koro	1
Mon	ɲah	2
Khmer	bokkal	L3

65. **rain**

Nihali	barsa:do	L
Bonda	mándo	L2
Sora	sipuṭap	1
Mundari	(daʔ)gama	2
Korku	ghama'	2
Mon	bray	4
Khmer	bhliəŋ	5

66. **red**

Nihali	==	0
Bonda	siye	1
Sora	ʒʔe: borreŋ	2
Mundari	araʔ	3
Korku	rata	L
Mon	bəkat	4
Khmer	krəha:m	5

67. **road**

Nihali	dáy	1
Bonda	luen	2
Sora	taŋór	3
Mundari	hora	4
Korku	kora	4
Mon	gloŋ	5
Khmer	t<hən>ol	6

68. **root**

Nihali	jaḍi	L
Bonda	regi	1
Sora	baléŋ	2
Mundari	red	3
Korku	ʒaṛi	L
Mon	ruih	1
Khmer	rɨk	1

69. **round**

Nihali	==	0
Bonda	==	0
Sora	yar-yár	1
Mundari	gol	L
Korku	==	0
Mon	khḍom	2
Khmer	mul	3

70. **sand**

Nihali	biṭil	1
Bonda	==	0
Sora	lakiʒ	2
Mundari	gitil	1
Korku	bitil	1
Mon	biti:	1
Khmer	khəsac	2

71. **say**

Nihali	==	0
Bonda	ḍɔ(ʔ)	1
Sora	su~- ga:m- ópuŋ	2
Mundari	men	3
Korku	men, mhen	3
Mon	gah	4
Khmer	tha:	5

72. **see**

Nihali	ara:y	1
Bonda	ʒu-	2
Sora	giʒ-	3
Mundari	lel	4
Korku	doʔ	5
Mon	ɲa:t	6
Khmer	ghaeɲ	7

73. **seed**

Nihali	==	0
Bonda	liki	1
Sora	abáy	2
Mundari	hita	3
Korku	==	0
Mon	==	0
Khmer	grap	4

74. **sit**

Nihali	péte	1
Bonda	leʔ-	2
Sora	gob-	3
Mundari	dup	4
Korku	suban, subaŋ	5
Mon	dʒoʔ	6
Khmer	ʔəŋguj	7

75. **skin**

Nihali	tól	1
Bonda	sali	
	úsa	2
Sora	usá:l	2
Mundari	ur	3
Korku	==	0
Mon	sna:m	4
Khmer	səpɛ:k	5

76. **sleep**

Nihali	ka:ppo	1
Bonda	lemɔʔ-	2
Sora	dib-mád-da-	3
Mundari	duṛum	4
Korku	kitiʒ	5
	taber	
Mon	stik	5
Khmer	ṭə:k	5

77. **small**

Nihali	base	1
Bonda	ḍau	2
Sora	sánna:	
	áʒidtíd', ánritíd	3
Mundari	huṛiŋ, huḍiŋ	4
Korku	sani	L
Mon	ɲi	5
Khmer	guc	6

78. **smoke**

Nihali	dhura:	L
Bonda	mugsiŋ, məksiŋ	1
Sora	umúd	2
Mundari	sukul	3
Korku	dhu~a	L
Mon	yak	4
Khmer	phəsɛ:ŋ	5

79. **stand**

Nihali	cippo	1
Bonda	tɔŋ	2
Sora	tánaŋ-	2
Mundari	tiŋgu	2
Korku	teŋgin(e)	2
Mon	tuiw	3
Khmer	ʒhor	4

80. **star**

Nihali	==	0
Bonda	mɔrtək, məmɔrtək	1
Sora	tui-túʒ	2
Mundari	ipil	3
Korku	==	0
Mon	saman	4
Khmer	phəkay	5

81. **stone**

Nihali	kurup	1
Bonda	bire, berɔ	2
Sora	áre:ŋ	3
Mundari	diri	4
Korku	gota	5
	dhega	
Mon	tmoʔ	6
Khmer	thəmo:	6

82. **sun**

Nihali	dewta:	L
Bonda	si`ʔi	1
Sora	uyuŋ	2
Mundari	siŋgi	1
Korku	==	0
Mon	==	0
Khmer	thəŋaj	3

83. **swim**

Nihali	==	0
Bonda	pɔ~ɔ~r, pɔŋɔr	1
Sora	é:ŋ-da:	2
Mundari	oyar	3
Korku	veer	4
Mon	biŋ	5
Khmer	hel	6

84. **tail**

Nihali	págo	1
Bonda	pəraʔ	2
Sora	əlá:	3
Mundari	cadlom	4
Korku	cud	4
Mon	ḍut	L
Khmer	kənduj	5

85. **that**

Nihali	enga	1
Bonda	gitin	2
Sora	únte	
	kun-	3
Mundari	haniʔ	4
Korku	dee	5
Mon	ʔigah	6
Khmer	naoh	7

86. **this**

Nihali	ere, he-re	1
Bonda	kɔʔn	2
Sora	kan	2
Mundari	niiʔ	3
Korku	==	0
Mon	woʔ	5
Khmer	nə:h	6

87. **thou**

Nihali	nye	1
Bonda	nɔ	1
Sora	ámən	2
Mundari	am	2
Korku	am	2
Mon	beh	3
Khmer	ʔəmbhao	4

88. **tongue**

Nihali	laiŋ	1
Bonda	leʔaŋ	1
Sora	lá:ŋ	1
Mundari	leʔ	1
Korku	lan, laŋ	1
Mon	kata:k	3
Khmer	ʔəṇta:k	3

89. **tooth**

Nihali	ménge	1
Bonda	gine	2
Sora	ʒʔi:	3
Mundari	ḍaṭa	L
Korku	tiṛiɲ, tiṛiŋ	4
Mon	ŋek	5
Khmer	d<həm>ə:ɲ	5

90. **tree**

Nihali	aḍḍo	1
Bonda	semuk	2
Sora	əne:b	
	ará:	3
Mundari	daru	4
Korku	siɲ	5
Mon	chu	2
Khmer	ʒhaə	2

92. two

Nihali	ira:r	1
Bonda	mbaʔr	1
Sora	bá:gu	1
Mundari	bar	1
Korku	bari	1
Mon	ba:	1
Khmer	bi:r	1

92. walk

Nihali	er	1
Bonda	uriŋ-	2
Sora	ɲa:-	3
Mundari	sen	4
Korku	sen(e)	4
Mon	twa:k	5
Khmer	ṭaər	6

93. warm (as water)

Nihali	==	0
Bonda	==	0
Sora	paɲúm	1
Mundari	lolo	2
Korku	cata	3
Mon	guim	3
Khmer	kədaw	4

94. water

Nihali	joppo	1
Bonda	ḍak	2
Sora	dʔa:	2
Mundari	daʔ	2
Korku	daʔ	2
Mon	ḍa:k	2
Khmer	dɨk	2

95. we (excl.)

Nihali	manê	1
Bonda	nay	1
Sora	əl-len	2
Mundari	ale	2
Korku	ale	2
Mon	puiy	3
Khmer	jaəŋ	4

96. what

Nihali	na:na:	1
Bonda	ma	2
Sora	ité:n	3
Mundari	cana, cina, kana	4
Korku	cuttha	5
Mon	nu	1
Khmer	səʔi:	6

97. what

Nihali	pa:ndher	L
Bonda	tuli	1
Sora	palú:	2
Mundari	puṇḍi	L
Korku	pulum	3
Mon	pata:ŋ	4
Khmer	so:	5

98. who

Nihali	na:ni	1
Bonda	ʒa	2
Sora	bó:te	3
Mundari	okoy	4
Korku	ʒee toneʒ	5
Mon	==	0
Khmer	ʔanakna	6

99. **woman**

Nihali	ba:i	L
Bonda	kuniʔram	1
Sora	ənsəló:	2
Mundari	kuṛi	3
Korku	==	0
Mon	brau	4
Khmer	strɛ:j	L2

100. **yellow**

Nihali	==	0
Bonda	saŋ saŋ	1
Remo	saŋ-saŋ-	1
Sora	saŋ-sa:ŋ	1
Mundari	saŋ saŋ	
Korku	pevaṛi	
Mon	ḍaʔmit	2
Khmer	lɨəŋ	3

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Some Remarks on John Bengtson's Comparison of Ainu and Nihali

Alexander Vovin
University of Hawai'i at Manoa

I have read with great interest John Bengtson's comparison of Ainu and Nihali. While I am in no position to comment on the Nihali data, as an Ainologist and general historical linguist, I can offer some thoughts on how Ainu data are handled and on the general prospects of the proposed Ainu-Nihali genetic relationship.

First, Bengtson has chosen his Ainu data very carefully. Several (non-significant) mistakes that were present in his draft version have been taken care of since, and the Ainu material looks good at the present time. On a note of caution, though, I have to add that he sometimes cites entries from John Batchelor's *Ainu-English-Japanese Dictionary* (1938) rather than Proto-Ainu reconstructions. Caution must be taken here, since Batchelor's dictionary has a very eclectic nature (thus, it does not differentiate between the colloquial and epic language, it does not specify dialects, etc.). Besides, there are occasional mistakes in this dictionary, as pointed out by Chiri Mashiho in his *Ainu go nyuumon* [*Introduction to the Ainu Language*] (Sapporo: Hokkaidoo shuppan kiga sentaa, 1966), so a word appearing *only* in Batchelor's dictionary and nowhere else is a dubious candidate for long-range comparisons.

Second, a number of comparisons proposed by Bengtson do *look* impressive. I recollect another attempt to relate Nihali to something else, that time to the so-called "Sino-Caucasian" macrofamily (I personally do not support the idea of this macrofamily), by N. Gurov in 1989, but it was not persuasive at all. Therefore, it may be that Bengtson is on the right track (with emphasis on the words *may be*). However, a word of caution must be exercised here as well, since so far Bengtson has presented a list of impressive look-alikes, which he has yet to prove to be genuine cognates, demonstrating the existence of regular phonetic correspondences. I did not check his lists in all details, and my knowledge of both Nihali historical phonology and morphology is zero, but it *looks as if* there might be a couple of interesting regular correspondences in the consonantism. I am less optimistic about the vocalic correspondences, but we will see what Bengtson will be able to come up with.

Third, there have been several claims recently regarding the genetic relationship of Ainu with Austronesian (Murayama Shichiroo 1992b, 1993, 1995), resurrecting the almost forgotten ideas of Leo Sternberg (1933); with Austric (Vovin 1993), originally suggested by Olaf Gjerdman (1926), as well as with Austric as a whole (Bengtson & Blažek 1996). Since Bengtson himself supports the idea that Ainu is an Austric language, it is not enough to demonstrate that Ainu and Nihali are related, it is also necessary to show that Nihali is related as well to other Austric groups, such as Austronesian and Austroasiatic (I do not share with Bengtson his opinion that Kadai is an Austric language too — I believe that it is totally unrelated). Unless such a procedure is followed, it will be very easy for Bengtson to fall for something like the Dene-Finnish macrofamily proposed by Morris Swadesh.

In sum, John Bengtson has presented interesting data, which may turn out to be significant. Now it is necessary to prove the suggested Ainu-Nihali relationship by conventional methods.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

On Nihali

Norman Zide
University of Chicago

The interest in Nihali in certain narrow academic quarters, lies in the argument that it is not a language at all but a 'so-called' or seeming language, and/or that it is a mystery, a lost — possibly 'Paleolithic' — language (something like the Tasaday of the Philippines, what Tasaday was purported to be but without the heavy public relations flak that surrounded it). It is, perhaps, the only remnant of an ancient — pre-Munda, pre-Dravidian, pre-Indo-Aryan — language family, with no living relatives, but perhaps a sister language of the language the Bhils spoke before they lost their own language and it was supplanted by the various Indo-Aryan 'Bhilis'. Nihali has been noticed by historical linguists for the very high percentage of borrowed vocabulary, and the variety of (proposed) sources for that borrowing, and the 'suspiciously simplified' syntax of the language. What is a mystery academically and popularly can be an administrative headache. What early notices of the Nihals we have describe them as nuisances, hill marauders and plunderers, 'caterans' who were 'incorrigible', and needed to be exterminated, and almost were on a couple of occasions.

I use Mundlay's spelling, Nihali, which represents the local pronunciation; Kuiper and others write 'Nahali'. Berger's paper goes into the history of the name. The name the Nihals use for themselves is Kalṭo or Kalṭu. It is due to the work of Professor F. B. J. Kuiper that Nihali has been brought to the attention of Indologists, and what we say here addresses matters that Kuiper has been the first to foreground, and to treat in impressive detail.

Nihali has been referred to several times as a 'so-called (sogennante)' language or something similar by Koppers; Konow, Kuiper, and even Fuchs expressed doubts about the language. The new many-volumned epitome of gazetteers and tribes-and-castes compendia, *People of India* (being issued by the Anthropological Survey of India), in its ninth volume, *Languages and Scripts*, recognizes and, on the strength of its own investigations?, finds the 'Nahals' as speaking — in different regions — Nimari or Korku (but not Bhili), but there is no mention of a spoken Nihali language. What is defective or 'so-called' about Nihali? Why is it not just a language, comme les autres? For Kuiper it is an argot (of what? or in what multilingual package?), and he talks of gaunersprachen (secret languages used by criminals); Koppers seems to doubt that it is a full (complete) native language, the first language of anyone, the assumption being that all Nihali-speaking Nihals (a small minority of those identified in official records as Nihals) are bilingual; their other — full — language presumably being the North Munda language Korku. Kuiper reports Koppers' mention of a collection of texts collected by Koppers and Fuchs, but Fuchs in a recent book, where he presents a considerable amount of information about the Nihals and says something about their language, makes no mention of any text collection. (Mundlay found and worked with Nihali monolinguals as well as bilinguals, and did collect texts. It is, as of the sixties and perhaps still, a first language, a home language, and most likely a 'full language', however we choose to define such a phrase. We will come to that later.) Konow was responsible for the data on the Munda languages and Nihali (in volume 4 of the *Linguistic Survey of India* [LSI]), and Koppers quotes his 1908 article: (that there is) '...one tribe, the so-called Nahals of Nimar, who were stated to speak Kurku in 1870, but who now speak (c 1908) a mixture of Munda, Dravidian and Aryan dialects', presumably our 'Nihali', or some variant of it. The evidence for their speaking Korku and nothing else in 1870, in any case, is shaky. The progress from monolingual Korku to some 'Nihali' seems unlikely. 'Nihali' has been in and out of the roster of Indian languages several times. Now you see it, now you don't.

Discussions of Nihali presuppose conjectural histories of the 'language', so that, for instance, it is not clear that calling it an 'argot' refers to present day use of Nihali usage (if Mundlay's data on monolinguals are accurate, and I think they are, then at least for some group(s) of Nihals Kuiper is mistaken) or to some earlier stage ('argotization'??) in the formation of Nihali. Is an argot 'stage' recognized in the formation of some component of other ('full') languages? If Kuiper is referring to phonological deformation and 'mutilation' — and he does talk of mutilation — does 'argot' have something to do with 'pidgin'?

Reasons adduced for doubting that Nihali — if there is one or a set of closely related dialects that are being consistently referred to in these publications — is a full-fledged language, and the first and /or only language of any speaker are the following: 1) the common association of Nihali-speaking Nihals (the estimate as of 1963 of the Nihal population was c 25,000, of which [Mundlay's estimate] perhaps ten percent spoke Nihali) with Korkus and Korku villages. Presumably elsewhere in Nihal territory (i.e., in de Candolle's Zones 2 and 3) they did not speak or know Nihali. (I doubt if anyone has made a careful investigation of Nihali language competence and use throughout the area. Mundlay has useful information of the Nihals she surveyed in the Melghat region.) Korku villagers I worked with when asked about the Nihali language (Mundlay assured me that there were Nihali-speaking Nihals living in the village) told me that the Nihals had no language of their own; they spoke Korku. The few extended descriptions of the Nihals are in books (Fuchs [1988], Hermanns, Koppers) primarily concerned with other groups, the Korkus or the Bhils, and this seems to be the characteristic angle of encountering and viewing the Nihals, when seeing them as anything but a source of civil disturbance and disruption. The few exceptions, papers addressing primarily Nihali matters, include the much quoted brief piece — the one source known to administrators or scholars having to find out something about the Nihals — in the *Tribes and Castes of the Central Provinces* (1916), and de Candolle's paper). 2) The apparent secrecy about the language and the general ignorance of its existence suggest to some that it isn't a language, but an argot or jargon used for certain limited purposes, the real language of these people (in Melghat/Nimar) being Korku, or in other areas — is the argot there completely gone? — some form of Nimari — or Bhili or Hindi or Marathi. 3) More interesting is the judgment that the language is limited, defective, perhaps a broken down descendant of an earlier 'full language', or a mixed language, and not adequate to the usual needs of linguistic communication. Thus the need for Korku, or some other 'full' language. 4) An examination of the structure of the lexicon has led Kuiper to suggest that in fact Nihali is an argot (see below).

But the information on which all these generalizations have been made is limited. Shafer and Kuiper independently exhumed Nihali from the brief description in the *Linguistic Survey of India* and observed that it was not — as Grierson/Konow implied (although not without reservations) — a North Munda language, probably closely related to Korku, but something else. The LSI has a few pages on Nihali. Bhattacharya on one short field trip collected a small amount of material on Nihali (see his article), but the further field trips to the Melghat area he intended (personal communication, S. Bhattacharya c 1966) were not allowed by his superior, the then Director of the Anthropological Survey, Nirmal Kumar Bose. Kuiper's thorough study of Nihali of 1962 was based on the Nihali materials of the LSI and Bhattacharya. That's all there was. All the surmises about the status of Nihali can be shot down by new and better observational data, and Mundlay provides some of that.

Nihals and Bhils, Nihali and 'Bhili'. For the connection of the Nihals with the Bhils, see Koppers (1948), and also Kuiper (1962) and Fuchs (1988). Koppers quotes Campbell (1880) who wrote that the Nihals 'are the most savage of the Bhils', but this, Koppers says, is not to be taken literally. The Nihal problem is complicated. (Koppers has a few notes about the Nihals in this book and in his Geheimnisse.) There is an extensive ethnographic literature on the numerous Bhil groups, and something, but much less, on the language(s). The 'Bhili language' is, apparently, a number of Indo-Aryan dialects of the regional languages in the extensive area of Bhil settlement (Gujarat, Rajasthan, Maharashtra, Madhya Pradesh — see Koppers' map). In some regions, Nihals have been long associated with Bhils, have lived with them and still do (see de Candolle). (There are no — or few? — Nihali villages; Nihals live in sections of Korku, Bhil, and other [which?] villages.) Fuchs also mentions Bhili-speaking Nihals as well as Korku-speaking and Nihali-speaking Nihals. The pre-Indo-Aryan language the Bhils probably spoke is lost, and we don't know its genetic affiliation. West central India is almost entirely Indo-Aryan speaking now. Presumably other linguistic families were more strongly represented in these areas in earlier times. There are other — fairly large — groups in central India, the Baiga for one, who now speak a variety of the local 'Hindi dialect', but who probably had their own, non-Indo-Aryan, language earlier. That 'Old Bhili' was related to 'Old Nihali' — that there was an ancient Nihali-Bhili family — is a plausible surmise (this was suggested by Koppers and by Shafer and accepted by a number of others), but as yet there is no linguistic evidence for it, and I have seen no strong claims based on ethnographic materials to support the case. I examined one lexicon of Bhili — Thompson (1895) — and found no vocabulary cognate with the Old Nihali vocabulary identified by Kuiper, Shafer, and myself in the data Kuiper used and in Mundlay's data.

Dravidian. Tibeto-Burman. Most of the Dravidian cognates adduced by Kuiper, Shafer, and Bhattacharya seem plausible. Pinnow in his review of Kuiper summarized his (Kuiper's) material on Dravidian influence: there are four strata (schichten) — (Kuiper speaks only of sources of Dravidian words, not strata) with c 47-50 examples (9 per cent of the total). If borrowing from Kurukh — one of the four strata — is relevant to the history of Nihali, it may be possible to date (approximately) some of the borrowing, given a hypothesis about the location of the Nihals at the time of known movements north of the Kurukhs (Oraons). Burrow wrote a short notice of the book, but had no comments on the Dravidian material. The review provides a concise description of Kuiper's intentions, materials, conclusions; the only comment he allowed himself was to remark Kuiper's 'considerable reserve' — resistance —

to accepting Nihali — the lexical remains after the borrowings have been extracted — as 'a language which is in origin quite independent' (of Indo-Aryan, Dravidian, Munda and Austroasiatic, Tibeto-Burman), which to Burrow 'seems most likely', and which, as I too see it, Kuiper's monograph makes a good case for. One would like a Dravidianist to go over the entire corpus (i.e., including Mundlay's material) and comment on the whole picture. I am no Tibeto-Burmanist but the Tibeto-Burman (possible) cognates — of Konow and Shafer and Kuiper — I find less convincing. Here too, much new material — data and analyses — is now available, and it would clarify several of the issues if someone familiar with it and the other Nihali-related material were to reevaluate the Tibeto-Burman connection. Certainly a few of the forms (Nihali *sunum* 'oil', from Korku and North Munda *sunum*) seem to have related forms in Tibeto-Burman, and there are certainly old Proto-Munda (and Proto-Austroasiatic) loans — whichever way the borrowing went (e.g., PM **kuXla* 'tiger'), but, for instance, the geographical information Kuiper provides on the proximity of a Tibeto-Burman-speaking group, Limbu — 'not greater than about 130 miles', to (present day) Santal groups may not be relevant to Nihali, although it is to North Munda since there is no reason to think that the North Munda connection (borrowing or whatever) was not primary, and that Nihali borrowed the form — as it did so many — later, from Korku. Kuiper mentions Konow's views on 'complex pronominalized Himalayan languages', and a Munda substratum as a contributor to their formation. Konow's views on these languages — and the Munda substratum — are not accepted by Tibeto-Burmanists today. Kuiper finds the Tibeto-Burman connection to be 'the most puzzling problem' of Nihali contacts, but, with caveats, goes on to find grammatical morphemes in Himalayan languages as (possibly) connected with Nihali morphemes.

Austroasiatic (apart from Munda). This discussion of the possible connections of Nihali with Austroasiatic is based on material in Kuiper (1962) and on various papers of Pinnow's, which are also taken up in some detail by Kuiper (1972). The linguists who have done considerable work on Austroasiatic (primarily Mon-Khmer) in the last thirty-five years or so, e.g., Shorto, Diffloth, and Ferlus have had nothing to say about Nihali, probably because they don't find it to be (interestingly or at all) Austroasiatic. The identification of West (Munda) and East (Mon-Khmer) Austroasiatic cognates in general (of course some languages have undergone more obscuring sound changes) has not been difficult. That the establishing of plausible Nihali cognates — the paucity of data making things that much harder — has been difficult and uncertain could be a result of several states of affairs, one being remoteness of relationship. Pinnow (1963) proposes a provisional ('the present state of investigation of the position of Nahali does not permit any definite judgment.') Western group of Austroasiatic languages which he calls Nahali-Munda, Nahali (now definitely judged not to be Munda) being Western N-M, and Munda being eastern N-M. He writes that the classification of Nahali is 'particularly difficult', in large part due to lack of data. The more interesting and difficult to explain connection of Nihali with Munda has to do with its morphology ('Its morphological system ...is obviously connected with that of the Munda languages.' (See details in 1966a, and some remarks on Nihali verb morphology and its implications below under 'What Kind of a [Contact-Shaped] Language is Nahali?'). Kuiper quotes Pinnow's 1963 conclusion 'We may perhaps come closest to the truth if we assume that Nahali possesses an isolated non-Austroasiatic stratum that has been partially replaced by an Austroasiatic stratum which has also provided Nahali with its inflection.' I would ask why 'Austroasiatic' here should not be replaced by 'Munda'. In 1965 in the Austroasiatic pronoun paper, Pinnow writes (again I am quoting Kuiper [1972]) 'the personal pronouns of the disputed language Nahali can be classified with those of the Austroasiatic family, even though they are rather markedly distinguished from the personal pronouns of the other groups.' The Nihali pronouns don't look like Munda pronouns, and Pinnow finds a few similarities of individual Nihali pronouns with forms of similar meaning in Austroasiatic languages, e.g., Khasi. I don't find these miscellaneous similarities indicative of genetic relationship, and Pinnow himself expresses doubts in the paper. But in his 1966b review of Kuiper, Pinnow finds himself increasingly persuaded of Nihali's fundamental Austroasiatic character ('Der grundlegend austroasiatische Charakter des Nahali schält sich so nach und nach immer mehr heraus'). Kuiper writes that 'my provisional attempt at an analysis of the case-endings and the pronouns did not confirm this assumption of an Austroasiatic provenance'. I agree with Kuiper in finding little evidence of Austroasiatic provenance. Kuiper's 'central problem' in 1962, 'how we must conceive the relations between that oldest Austro-Asiatic stratum and the other unidentified component of the language' should perhaps now be decentered.

Munda. Apart from the numerous borrowings from Korku, what has Nihali borrowed from Munda, or Munda from Nihali? First, of course there is no assurance that all the Korku borrowings have been identified. And in the absence of sufficient possibly cognate vocabulary, no setting up of sound correspondences (Nihali-Kherwarian, or Nihali-South Munda) is possible, so that one goes by one's own intuitions about relations of words — in one's own style of negative capability. Examining the sets of words on Kuiper's page 39, 'A. More closely connected with North Munda (Kherwari)', and 'B. More closely connected with Central and South Munda', I find several of the seven items in A. unacceptable or implausible, most importantly, *ṭe-* (Mundlay *té-*) 'to eat', which does belong here, but in set B. Of the items in B. the word for 'father', *a-ba*, is pan-Munda, reconstructable — and not a loan — for Proto-Munda. The most interesting — and to my eye the most solid — forms are *be-* 'to give' (Mundlay *be-*), **er, ier-* 'to go', *piy-* to come', and *ṭe-*. Although we have only these four words, the connection here is more persuasive to me than anything in Kuiper's Munda alignments and the claims of cognation that go with them. These four do have good parallel forms in one or another branch of South Munda, and apparently no related forms in North Munda. I don't accept Kuiper's Santali *ạtiñ* as likely to be connected with Nihali *te-*. (South Munda [SM] branches into Kharia-Juang [KJ] and Koraput Munda [KM]. Koraput Munda branches into Gutob-Remo-Gtaʔ [GRG] and Sora-Juray-Gorum [SJG]). The actual forms and their antiquity — subfamily membership — will be discussed elsewhere. We give here rough reconstructions: GRG **bɛd-* 'to give', GRG **pɨŋ-* 'to come', KM **tej-* 'to serve food', KM **ɨr/er-* 'to run, jump, move'.) What do we make of this? The critical question — crucial to a hypothesis of South Munda subfamilies having borrowed from Nihali words that have no congeners in Mon-Khmer — of Austroasiatic cognates for these SM forms is as yet unanswered. That Nihali could have come into contact with South Munda languages is not at all unlikely. Certainly, some of the Koraput Munda languages (e.g., Gutob Gadba, now spoken only in Koraput District, Orissa was spoken further east, in Bastar, ca. seventy-five years ago, if not more recently), and we have no realistic notions of where and how the Nihals earlier ranged or came from. What sort of contact situations between what sort of groups, Munda and Nihali, could have resulted in the borrowing of basic lexical items? If the power of the Nihals earlier (as it was some time later) was military, their success as marauders (like the [SM] Remo [Boṇḍa] today but in a smaller way, contained by the local district administration — or like the Comanche and other Plains Indians) what sort of linguistic impression would we expect the marauders' language to make and under what conditions? Was there (intermittent?) occupation of the raided territories? Intermarriage? The claim that some of the SM languages (Juang, for one) spoke something else before they adopted the ancestor of the Munda language they now speak is not new. That 'something else' could have been Nihali, or a sister language of Nihali. There are many possible scenarios to account for the lexical similarities (borrowing, presumably), but I want to affirm the importance of the identification (by Shafer and Kuiper) of these forms; they are less questionable and (therefore) more important than the other putative linkages proposed. In Kuiper's discussion of Nihali and Austroasiatic, he writes '.. the circumstance that the non-Kurku elements of the Nahali vocabulary cannot be attributed to any one of the subgroups would seem to point to the conclusion that the older Munda stratum in Nahali stands somewhat apart from the sub-groups into which Munda is divided. Berger arrived at the same conclusion.' If the Kherwarian similarities can be discounted, and I think they can, and the A and B sets are revised and realigned as proposed above, then perhaps (the corpus is still too small, but we can perhaps find more forms supporting this hypothesis) it is precisely one subgroup, South Munda, or perhaps some sub-family or subfamilies of South Munda that show(s) these lexical relationships, and it is South Munda (the SM family and/or one or another of its subfamilies) that has the connection with Nihali. I suggest that SM or KM has borrowed from Nihali, and Kuiper's and Berger's conclusion is wrong.

Argot. In 1962, Kuiper writes 'In the case of Nahali, it is true, there are no certain indications of an analogous origin (he has been talking in the previous paragraph of metonymy and mutilation in speech disguise in various secret languages of the subcontinent) of the names of parts of the body, etc., which categories are also in Nahali etymologically unexplained. Still, it may be useful not to forget that some of the obscure Nahali words may also belong to an argot, and need not necessarily date back to a linguistic pre-history of India.' This is an interesting and useful warning. He mentions *jiki* 'eye' as perhaps connected with Santali *jhiki miki, jiki miki* 'splendid, resplendent, shining, radiant' — and also notes Ainu *shik(i)*. (Kuiper in his earlier work on Proto-Munda words in Sanskrit brings in echo pairs of this sort, none of which are, as yet, reconstructable for Proto-Munda). In his 1972 review, he writes (in a discussion of my inadequate treatment of Nihali in a survey article on the Austroasiatic languages of India): (Zide's) 'observations contain nothing new except the confirmation that Nihali is actually an argot, as had been suggested (by Kuiper) in 1962'. I did not think or say that Nihali was an argot, but that it was likely that 'Nihali was used as a more or less secret language'; Navajo was used during World War II by the United States military as a 'secret language' because it was unintelligible to the enemy, this did not make it an argot. It seems possible that some of the obscure Nihali words may belong to an argot, but there are — as Kuiper shows other reasons for obscurity. Despite the interesting and not irrelevant discussion of gaunersprachen I see no good reason (the *jhiki miki* forms don't convince me) to claim that Nihali is an argot (now? at which previous stage? all of it? some section of the vocabulary? which?). It may be that the phonological distortions, lexical substitutions (rhyming slang, etc.) found in (other) secret languages are responsible for some of the Nihali vocabulary — certainly 'some of the obscure Nihali words' may be argot, i.e., the result of speech disguising transformation and substitution, but this is something suggested here, and in no way demonstrated, and if it was something like rhyming slang (as in Cockney) there would be no way of retrieving the baseforms, and thus of proving that there was, in fact, this sort of distortion. Kuiper' s reasons for proposing his argot hypothesis seem to be, first, the social position and criminal activities of the Nihals (which don't guarantee their possession and use of an argot), and, second, certain speculations about a few words in the old Nihali lexicon. I find the case unproven. I learn from Hal Fleming about 'jargons' in small, low status hunting and gathering groups in East Africa where a small stock (c 40 words) of 'jargon' has been recorded and the casual conclusion drawn from this short vocabulary is that the language 'is a jargon'. One needs to see how much and what segment of the lexicon is (speech disguise-derived) 'jargon' and what else 'the language consists of'. Kuiper's case for an 'argot', more explicitly, is the following (1962, pp 11-16): he first takes up the low status of the Nihals as a 'despised social group' and notes that other low status groups in India have secret languages. He then introduces various kinds of phonological 'mutilation' found in such secret languages. He notes that words for body parts are commonly replaced in secret languages by disguised forms, and goes on to gives the sources and derivations of some of the replacement forms, and, a bit later, suggests that Nihali *jiki*, 'eye' might perhaps (originally) be a descriptive term. All of this is suggestive, but hardly probative, and I don't find it persuasive. In 1972, he is more positive about the argot hypothesis, and adduces some material (e.g., on Vedda) that might suggest analogues for what happened to Nihali, but again with nothing closer to a proof.

The quest for Nihali seems in some subsubtext to reveal a plot, one that Professor Kuiper most probably did not intend and would not accept: the voyage to Nihali as the grand occasion for wide-ranging exploration of Indo-Aryan, Dravidian, Austroasiatic, and Tibeto-Burman. Then, the winnowing of the Nihali lexicon, and the extraction of contact-derived matter. What is left is a small cache of semi-precious Old Nihali words, but this does not satisfy . A second voyage, on the Argo(t) — the golden fleece was plastic after all — leads to the discounting and discarding throwing overboard — of some of that old Nihali vocabulary. Apart from these substantial, scholarly souvenirs de voyage — what is left of 'Nihali'? More, I think, than the argot hypothesis seems to allow. Pinnow (1966b) agrees with Kuiper on the importance of the argot hypothesis ('Der Hinweis auf den möglichen Argot-Charakter des Nahali ist eins der wichtigsten Ergebnisse der Arbeit Kuipers, und seine Bedeutung kann nicht genug unterstrichen werden.')

We need to distinguish a 'functional argot' — i.e., the use of a (secret) language for concealment, from an 'argot' (secret language) formed through processes of word-deformation, substitution, etc. (see Guiraud and Mehrotra. Mehrotra isn't aware of Kuiper's monograph. Kuiper is not aware of some of the earlier material mentioned in Mehrotra, e.g., Sleeman on the language of the Thugs. The two discussions and bibliographies taken together provide a good survey of secret languages in India through c 1966.)

The parallels with Sri Lankan Vedda and Rodiya are interesting (Vedda — Kuiper, quoting de Silva — 'is a creole based on an older Vedda language with Sinhalese as the second contributing factor', (Kuiper) 'Rodiya is a secret language in which non-Sinhalese items are used in Sinhalese structures'). Can (our) Nihali be a creole based on an older Nihali language? Individual factors and contexts may be shared by Nihali (under various conjectures) and Rodiya, Vedda, etc. But, as Kuiper's data show, none of these cases is closely parallel to the Nihali situation. Nihali exhibits a wide range of linguistic contacts, many more than were available to Vedda or Rodiya. What the time scale is — in any of these cases — is still unknown. For Nihali, we assume that there was considerable mobility in a fairly extensive multilingual territory, so that such partly similar contact situations as that of Brahui or Vedda or Rodiya with massive borrowing (or deformation?) but less extensive linguistic contacts are only partly similar. Perhaps the language of the Thugs should be more closely examined. The Romani (Gypsy) sociolinguistic situations — one or more of them — seem more like what the Nihali situation(s) may have been, but for Romani we know where the people came from and, roughly, when (it is relevant that earlier speculations about the Gypsies posited a much more ancient exodus than the one scholars later reconstructed), and we know their original — pre-exodus and prewandering — Indo-Aryan language and a fair amount about the languages they came in contact with, whereas for Nihali the ancestor language is presumably unattested outside the (obscure) Nihali lexical corpus, and some of its possible contacts — as proposed by Kuiper — have yet to be more firmly demonstrated. And as Kuiper has shown, the proportion of borrowed vocabulary in Nihali is very high, presumably much higher than in any of the Romani dialects (what the corresponding figures for Vedda are I can't say.) As with most everything else about the Nihals, we know little about their social or occupational history. They do not now and did not in the recent past own and cultivate fields (and there is no evidence that they practiced slash and burn agriculture) or cattle. In the Melghat, they seem to have been associated with the Korkus (themselves known earlier as freebooters, but now settled agriculturalists), but we don't know how far back the connection goes. They were probably hunters and gatherers, and did more and less raiding (including cattle raiding?) of neighboring sedentary communities. Fuchs mentions that they are skilled trainers of dogs, and this skill is appreciated by the Korkus. One wonders about their mobility as 'caterans'. Did they have horses? They don't now, and neither do the Korkus, although (see Fuchs) there are representations of horses commonly on Korku wooden funerary tablets, and a taboo on eating horse meat. Horses in that area would be expensive to keep, not particularly practical — bullocks are at least as efficient as ploughing and cart animals, and more docile, and healthier in that terrain. The word for 'horse' (see Kuiper) is *māv*, which Kuiper connects with Dravidian (e.g., Telugu *māvu*) and possibly Tibeto-Burman and Tai forms. The Korku and other Munda forms (old borrowings) are not related. Hermanns has origin stories from Nihali informants according to which in earlier times the Nihals and Korkus were one people. (Not surprisingly, Korku informants deny this.) This means, I assume, that they (all) then spoke a form of Nihali. Later, the Korku upgraded themselves (giving up beef, certain low occupations, etc.) and, presumably, adopted a North Munda language from high status invaders/immigrants to the region. The Nihals were downgraded, and the Korkus have maintained the social distance (see Fuchs for Korku-Nihali interactions). There are references to Korku presence and activity in the fifteenth century, and to the Nihals 'at the time of Akbar', i.e., the latter part of the sixteenth century, in both cases as hill robbers and freebooters. See Fuchs (1963) also on the antiquity of the Korkus (and, by implication, the Nihals who, these writers would claim, probably were in there earlier if in fact they were a different group) in the region, and that of the other North Munda (linguistic) groups in Bihar and adjacent regions. There are, of course, the usual putative identifications with peoples mentioned in the Ramayana (as *rākṣasas* — demons; see Fuchs [1988], but also Zide [1972]) on Khara.) The social position and the marginal occupations of the Nihals suggest that they may well have had and used a secret language, early and late. Information on this may still be obtainable. That this is an argot (or that they use an argot — and what relation that argot has with 'Nihali') has yet to be proved.

What kind of a (contact-shaped) language is Nihali? A much more thorough treatment of this topic is called for, but I offer here one possible scenario showing schematically how Nihali may have come to its present state. (The data on the Nihali verb can be found in the *Linguistic Survey of India*, Pinnow [1966a], Bhattacharya and Kuiper [1962] as well as in Mundlay and Lynch, sources that were not available to Kuiper but that don't describe a system that is significantly different.) An examination of what Kuiper meant and might mean by 'argot' might introduce various linguistic — sociolinguistic — historical scenarios and tentatively try to place various statements and implications of Kuiper's in such a scheme. Here is a preliminary attempt at doing that (all of the assumptions and stages are arguable):

1) the (Old) Nihali language — not Austroasiatic — was spoken (where? perhaps in west central India; when?) as a first language by a (perhaps nomadic) group, probably not agriculturalists, and probably not pastoralists either. These people may well have been bi- or multilingual. It was a representative of a family no longer found in India (apart from the words inherited from that older lexicon in modern Nihali); there may well have been related languages at earlier times.

2) In the course of wandering in or to the eastern parts of central India (I won't try to break this down into ordered stages), there was borrowing from South Munda, i.e., South Munda from Nihali (see the words discussed in the 'Munda' section) and probably from South Munda into Nihali (not that can we identify which is which with much assurance. The preglottalized finals -*'ḍ* in *bɛd*, and -*'j* in *tej*- are not reflected in Nihali. One could claim it is likelier that the Nihals lost the distinctions among the preglottalized consonants, and that this was likelier than that SM forms borrowed from Nihali created them. Perhaps.). The Nihals may have been more powerful at this time, more dominant — powerful as raiders — it is less likely that they were more technologically advanced in some ways than the South Munda groups (but not in agricultural techniques?). The morphosyntax of Nihali, whatever it

was — I am assuming not much of it is now left and/or identifiable — at this stage, whatever 'natural' changes it may have undergone, is fairly intact.

3. There was some borrowing from Dravidian, massive disruption of the Nihali community (or communities), perhaps the decimation of the community on the orders of local rulers and chiefs (see Fuchs [1988], or earlier traumatic reduction and breakup of the community (due to what? natural catastrophe, i.e., disease??). (Kuiper quotes Forsyth on 'the aboriginal races' (having been) compelled to retire to the mountains before the Hindu invaders ... A few remained in the country occupied by the Hindus, chiefly in the position of agricultural serfs, or watchers of the village', a description which, somewhat modified, could apply to the Nihals in the (non-Hindu) Korku villages. Historical speculation about the position of the Nihals, early and late, is constrained — stymied — by the lack of information on where 'they' were when.) The old morphosyntax breaks up. Probably the Nihals (some group of them, one that remains, later, 'Nihali-speaking') are bi- (or multilingual) and lean more heavily on one other language for many vital functions, but retain some older Nihali, e.g., at least some of the lexicon. (On these massacres: Fuchs [1988] writes that 'in the latter part of the eighteenth century the Mughal power was slowly supplanted by that of the Marathas who soon began to impose heavy taxation not only on the prosperous farmers in the plains but also on the tribals in the hills ... The tribal chiefs retaliated by long drawn-out guerrilla warfare against the invaders.... (this led to the) massacre ... of a whole tribe of Nahals, men women and children, by a body of Arabs in the service of Scindia.' According to Russell and Lal 'in the times of Akbar (late sixteenth century) ... the Raja of Jeetgurh and Mekote (reports an account of) a treacherous massacre ... of a whole tribe of these Nahals (presumably because their marauding went too far) ... in reward for which he got Jeetgurh in Jageer.' There is some question as to whether the term 'Nahal' is used for Nihali-speaking ancestors of the Nihals we are talking about, but it seems likely.)

4. The older language was remade and socially reconfigured — with a 'creole-style grammar', this perhaps for (some) use as a secret language, perhaps with argot-style speech deformation in some of the lexicon. The tense/aspect, etc. markers — most of the morphemes used — as Pinnow points out (1966a = Kuiper's 1960a) are familiar to the linguist from a number of other languages in the area, but they are not used in the same ways and have different meanings; this doesn't look like a case of 'natural genetic' inheritance. In the Gutob (SM) language (and probably, earlier, elsewhere in SM as well), the tense suffixes in the positive conjugation are identical with suffixes with unaccountably (so far) different meanings in the negative conjugation. How this developed is baffling. It makes the language a bit harder to learn, but this seems to be a 'genetically natural development', whatever it developed from, whereas the Nihali salad of verbal morphemes seems to show a disruption, a bad break — and to be something else. The language is heavily relexified, Melghat Korku being the lexifier. (Kuiper's figures on borrowing from Korku are misleading in that much — most? — of the borrowing from Indo-Aryan has been borrowing from Korku also, borrowing of words Korku borrowed from Indo-Aryan.) The latter assumption has implications about the antiquity and nature of Korku-Nihali connections. If, anciently, the Korkus and Nihals were one people, speaking an old form of Nihali, and then 'the Korkus', upgrading themselves, separated themselves and adopted a North Munda language, were the Nihals and Korkus still in contact in the kind of relationship (Nihals as inferiors of and servants of the Korkus) we find today, and which the literature tacitly suggests has been the case for some time? If such was the case, the Nihals would have been — as they are now — bilingual in Korku (and, perhaps, as now, familiar with other local languages as well). This would suggest that the borrowing from Korku started very early, and not, as it looks, fairly late. It is simpler — until someone can show that this could not have been the case — to say that while the Korku-Nihali connection may be old, the heavy relexification dates to the comparatively recent period when the Nihal community was badly broken up, reduced, scattered, and that then some of the Nihals retreated to the Korku area, and recovered a social organization and a Nihali language. If we ask why Nihali — some sort of Nihali — survived at all, perhaps Kuiper's argot hypothesis suggests part of an answer. As to possible influences of — direct contacts with — non-Munda Austroasiatic languages, Tibeto-Burman languages — as the above discussion indicates, I have my doubts about these — they can have been acquired in the wander years of the Nihals, and clearly the Nihals have moved around. This scenario rules out the 'fundamentally Austroasiatic character' hypothesis of Pinnow's. If Nihali has a fundamentally Austroasiatic character, and acquired it in the usual historical linguistic ways, then my scenario is all wrong — unless the connection is remote indeed, in which case we want to hear more from Pinnow about that fundamental Austroasiatic character.

Questions of syntax that have not been examined, or mentioned in this introduction, can be investigated more closely using Mundlay's material. In this 'remaking', was what was *not* borrowed (at one stage or another) and is retrievable and assignable to 'Old Nihali' just a small set of words? We have said nothing about Nihali syntax, and how it compares with Korku or Hindi or Marathi or (some) Dravidian. The SOV word order, the use of postpositions, etc., are common to all the languages in the area (I do not speak of Tibeto-Burman). But it is not clear, for example, what relative clause(-equivalents) are like in Nihali, the Indo-Aryan pattern(s) being different from the Dravidian and Munda patterns. Perhaps a closer examination of Nihali will show traces of earlier structures. How 'simplified' or 'reduced' (simplified or reduced from what)? or 'creole-like' is Nihali grammar? Pinnow (quoted by Kuiper, who had just noted the absence of Dravidian influence in the [morphology of] the verb) suggests that the verb system is like that of Proto-Munda. I don't see this.

Younger historians of the region (they need to know Marathi) might be encouraged to look at the Nihals, and their history and place in history. The currently familiar and congenial problems having to do with colonialism, peasant rebellions, kingship and the discoursing that goes with them probably won't take them very far with Nihal history, but that would be one more reason to pursue it. If concentrated subalternity is of interest, the Nihals are the subalterns' (Hindi- and Marathi-speakers') subalterns' (Korkus') subalterns. How 'other' in the (non-urban) Indian scene can one get? Perhaps some illumination of the linguistic problems will come out of a better understanding of where the Nihals were and what they were doing and saying to whom.

The paucity of data on Nihali has not prevented wider comparativists — 'long rangers' — from finding (tentative) homes for it. Thus, J. Bengtson (1994) finds a place for Nihali in his Macro-Australic superstock.

The most substantial section of this paper, 'What Kind of a (Contact-Shaped) Language is Nihali', is much reduced here. It grew too long, and needs to be longer. To take up some of the problems coming out of Kuiper's dense and thorough treatment of the Nihali material available to him and Mundlay's material as well will require more work, much more consideration of the morphology and syntax. It is to be hoped that Mundlay's sketch of Nihali grammar, and papers on the sociolinguistic situation of the Nihals will also be published in the near future. with more results of more research on the Nihals.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Comments on the Basque-Dene-Caucasian Comparisons

Sergei A. Starostin
Russian State University of the Humanities
University of Michigan, Ann Arbor

For a long time, I was hesitant about joining the discussion that started with R. L. Trask's anti-Basque-Dene-Caucasian paper. The main reason for this was that I find discussions of this sort generally not productive. When I myself meet this sort of criticism, I simply try to correct my mistakes and to take account of them in further work. However, having read the Editorial I in *Mother Tongue* 1, 1995, I decided to participate. The editors (basically, I believe, H. Fleming) started accusing the Muscovites (for some reason twice spelled as Muscovites and twice as Moscovites on the same page 229) of a "startling whiff of paranoia" and, using characteristic Communist terminology, of "having abandoned comrade Bengtson". I also learned with some interest that "one of the editors has been doubting the Moscovite reconstructions for several years now", and they both (H. Fleming and A. Bomhard) think that "using the proto-Caucasic reconstruction of the Moscovites as a central focus caused Bengtson a great deal of trouble". Well, I doubt many reconstructions myself and am constantly trying to improve them, and I would be extremely grateful to H. Fleming (and to everybody else) for positive suggestions — if any. It is certainly possible that our reconstruction did harm to J. Bengtson's Basque-Dene-Caucasian parallels; unfortunately, I do not know of any other North Caucasian reconstruction, and without it, Basque couldn't be compared to anything Caucasian, and the discussion would probably not even have started.

Before I get to the point, I would like to elucidate my position concerning Dene-Caucasian and long range comparison in general.

1. It is well known that linguistic relationship cannot be disproved (for some reason the most ardent opponents of long range comparison like to repeat this point over and over). It turns out, however, that linguistic relationship cannot be proved, either. It is always rather a matter of faith, or, to put it into more academic words, of consensus among specialists. Mathematics clashes against mathematics (because for most philologists, the strange mathematics employed by Don Ringe sounds scientific enough), comparative method runs into exhortations about chance coincidences, and positive arguments are encountered by flat refusal. It is therefore important just to go on working — of course, appreciating valid criticism, and, with some luck, waiting for common approval ("consensus") to come.

2. I believe that the Sino-Caucasian (and, more broadly, the Dene-Caucasian) hypothesis is valid, at least in what concerns the relationship between North Caucasian, Yenisseian, Sino-Tibetan, and, with some reservations, Na-Dene. I wish, however, to stress that the Sino-(Dene-)Caucasian case is different from Nostratic. Illich-Svitych was lucky because he compared well known linguistic families with existing etymological dictionaries (except for Altaic). Nothing of the kind was available for Sino-Caucasian when I first proposed the Yenisseian/Sino-Tibetan/Caucasian relationship in 1982. Since then I have been trying hard to improve the situation in each of the three mentioned fields, and I hope that now the overall picture is close to that of Nostratic in the early sixties. I only wish that comparative Na-Dene materials were published, too (by S. Nikolayev, or by his opponents), so that the supposed fourth branch of Dene-Caucasian could be also open to investigation. There is still very much to be done in each individual branch and the reconstruction of Proto-Sino-Caucasian is still very far from any perfection, but nothing so far has shaken my belief in the genetic relationship of the languages in question.

3. What about Basque, Burushaski, and Sumerian? All three are linguistic isolates, in the sense that none of the three languages has any close relatives. The case here is even more difficult than with Japanese — which, I think, is more closely related to Korean than to other Altaic languages. Before I proceed with evaluating the Basque pros and contras, I would like to state explicitly that I am not yet finally convinced of their inclusion in Dene-Caucasian, although I'm inclined to agree (more in the case of Basque and Burushaski, somewhat less in the case of Sumerian). The theoretical situation here is the following.

Statistical considerations tell us that to be considered related languages must have no less than 10% of their most basic vocabulary in common (within Swadesh's 100-wordlist). If they have less, they may also be related (just separated too long ago), but a figure of about 5% may also be due to a chance coincidence. To distinguish between these two cases, we must be able to compare intermediate reconstructions: if modern languages have a lesser percentage, but it grows while comparing their respective proto-languages, there's a high chance that the languages do belong to the same branch; otherwise the resemblance might be fortuitous. Now for isolated languages like Basque, Burushaski, and Sumerian, we are unable to provide intermediate reconstructions (except for very shallow ones), so the situation here is much more complicated than even with North Caucasian, Yenisseian and Sino-Tibetan.

John Bengtson, however, has made a courageous attempt to compare Basque with Dene-Caucasian, adducing several hundred possible cognates. Now we have to evaluate this evidence and judge whether there is something in it or it is completely worthless.

Bengtson has made one significant mistake — unfortunately, not at all uncommon among long rangers. He does not pay enough attention to the established tradition of Basque historical phonology and etymology, notably to the works of Michelena, Trask, and others. Now this is the one thing that specialists in any field never forgive. Small wonder that Bengtson's papers meet a violent opposition from the Vasconists' side.

Let us, however, see if the case for Basque-Dene-Caucasian is really as bad as Trask is trying to demonstrate.

I am not a specialist on Burushaski, Na-Dene, or Sumerian; I shall not therefore comment upon etymologies including only those languages, without North Caucasian or Yenisseian parallels (Trask for some reason ignores Sino-Tibetan, see his remark on p. 10). This takes 32 items out of 317 (## 5, 17, 29, 34, 43, 45, 56, 60, 62, 70, 71, 77, 82, 96, 119, 124, 126, 131, 136, 140, 147, 151, 156, 174, 207, 212, 230, 232, 242, 254, 260, 279). There are also etymologies without any Dene-Caucasian parallels (considered to be borrowed from Kartvelian, Berber, or Egyptian, and thus irrelevant to our discussion; see ## 22, 27, 63, 67, 181, 193, 204, 239, 285, 300, 308), and two cases without a Basque word (Spanish words considered to be borrowed from a Bascoid source: ## 50, 216). This all leaves us with 272 etymologies to discuss (actually, somewhat less, because several etymologies are listed twice in Trask's paper, with cross-references).

All etymologies fall, from my point of view, into several categories:

1. Faulty from the Basque side. I think Trask's criticism is completely justified for the following 29 items: 2 (**aberats**), 3 (**abets**), 6 (**aditu**), 9 (**ahal**), 16 (**alderatu**), 30 (**arroda**), 79 (**busti**), 90 (**eme**, added by Jacobsen), 97 (**eskubarne**), 109 (**garkotxe**), 111 (**gela**), 163 (**ilu**, proposed by Chirikba), 166 (**intzigar**), 194 (***kala**), 196 (**kasko**), 200 (**kuma** — cf., however, objections by Jacobsen), 208 (**lorratz**), 211 (**makutsik**), 215 (**mara-mara**), 217 (**matel**), 268 (**toska**), 280 (**uhain**), 286 (**urtxakur**), 287 (**uxuri**), 290 (**-xe**), 295 (***zaki**), 297 (**zamar**), 299 (***zaro**), 314 (**zuku**). In all of these cases, the Basque word is either a transparent Latin loan, or has been misglossed by J. Bengtson, or is a rather obvious compound.

2. Faulty from the Caucasian or Yenisseian side. I have to admit that such cases exist among J. Bengtson's (and V. Chirikba's) etymologies. I mean the following:

    (47) Bq **barrabil** 'testicle': PY ***bajbVl** 'kidney'. The Yenisseian form is most probably an old compound < ***baj-pɨˀɨl'** (**pɨˀɨl'** 'intestine'), with ***baj** corresponding rather to PST ***bhia** 'spleen' (see Starostin 1995: 206).

    (61) Bq **ber-** 'self': Abkh. **a-bri** 'this' (Chirikba's etymology, based on earlier propositions by Dumézil and Lafon). The Basque form, quoted as **beri** 'this same' by Chirikba, has been corrected by Trask. I agree that the semantic side is weak here, and I must add that the Abkhaz form is easily analyzed as **a-b-ri** (as opposed to **a-ri**, **a-ni**, **a-b-ni**, see NCED 321).

    (78) Lezg. **firi** 'mane' etc., compared by Bengtson with Bq **buru** 'head', go back to PEC ***ƛ̱wɨ̆rV** 'horn; braid, mane' (NCED 771-772), which Bengtson elsewhere (# 4) compares with Bq **adar** 'horn'.

(101) Arch. **hiliku** 'fly' is absolutely isolated and has a rather strange morphological structure. I cannot completely refute Bengtson's comparison with Bq **euli** 'fly', but I doubt it very much.

(127) PY ***čɨʔs** 'stone' should be rather compared with PEC ***c̣ăč̣wV** 'small stone' (see Starostin 1995:218) and thus should be kept apart from Bq **haitz** 'rock'. If the latter really goes back to **anitz* or **anetz*, it might be possible to compare it with PNC ***ħ̄ǝmVc̣ŏ** 'stone' (NCED 516), with a not uncommon variation between **-m-** and **-n-**.

(157) Bq **idi** 'ox': Abkh. **-tǝ** 'ram' etc. (Chirikba's etymology). The WC forms most probably go back to PNC ***dwănʔV** 'sheep, ram' (NCED 405-406), and thus should be kept apart from the Basque word.

(246) Bq **pipil** 'bud': PY ***bajbVl** 'bud'. This is a sad misunderstanding. The etymology belongs to M. Ruhlen (in Starostin & Ruhlen 1994). I must say that I knew about Ruhlen translating my Yenisseian etymologies, but I never saw the final text until it was published (and I do not bear responsibility for the Basque, Burushaski, and Na-Dene forms included by Ruhlen). Although in general the translation is OK, PY ***bajbVl** 'bud' is a mistranslation of Russian 'почка' which means (due to a quite fortuitous homonymy) both 'kidney' and 'bud'. The form actually means only 'kidney' and is of course the same word that Bengtson had compared with Bq **barrabil** 'testicle' (probably also incorrectly, see above).

(262) I don't quite see what PNC ***čɨnHV** 'game, animal' has to do with Bq **soin** 'body, torso'. I would rather think about comparing the Basque word with PEC ***s̳ɨ̄nŏ** 'long bone; limb, paw, part of body' (NCED 963-964).

(296) Bq **zakur** is an interesting "Wanderwort" (cf. also Kartvelian ***ʒ́aγl** and probably borrowed into Kartvelian from Proto-Nakh ***ǯʔali**), but it certainly has nothing to do with Budukh **sokul** 'fox' (not **cäkul**, as in Bengtson's rendering). The latter is a metathesized development from Proto-Lezghian ***s:oIl-Vk** /***s:oIl-Vḳ** with a diminutive suffix (cf. outside Lezghian forms like Lak **culḳ-**, Batsb. **coḳal** etc.), going back to PNC ***chwōlĕ** 'fox, jackal' (NCED 324-325).

(307) Tsakhur **ǯil** 'navel', to my knowledge, does not exist (the word means uniformly 'earth, ground' in all Tsakhur dialects), thus the comparison with Bq **zilbor** 'navel' is hardly justified. Recently, J. Bengtson proposed some other comparisons for the Basque word, but they must be discussed separately.

(317) PY ***tat(aŋ)** 'straight' is a reduplicated stem probably corresponding to PEC ***=ītV̄**...

I should note the relatively small number of faulty comparisons on the North Caucasian side of J. Bengtson's data, as opposed to the large number of mistakes in the Basque material. I should also have to remind Dr. Fleming (who complains about Bengtson's using our North Caucasian reconstruction) that the quality of North Caucasian material in the works of Lafon, Bouda, and other predecessors was simply abominable. Well, maybe I am misjudging their evidence...

3. A large number of comparisons which do not seem very impressive to me (and could be perhaps due to chance), but which Trask is not able to refute at all. Interestingly enough, he leaves several comparisons without any comments, sometimes stating only that other authors had similar propositions before (so what? does this fact automatically "destroy" the comparisons? I could not quite understand it). In many cases, the Basque word is considered to be of "Romance" origin, but the Romance words themselves could be borrowed from some non-Indo-European substratum, closely related to Basque. Such cases are too numerous to be discussed here, but they constitute the bulk of Bengtson's data.

4. A substantial amount of quite plausible, both phonetically and semantically, correspondences. These include:

(10) Bq **ahizpa** < ***an-iz-pa** 'sister (of a woman), with ***-iz-** corresponding to PNC ***=ɨ̆ć̱ĭ** 'brother, sister', PY ***b[i]s** 'brother, sister', PST ***čə̆jH** 'elder brother, sister'.

(14) Bq **akain** 'tick': PNC ***q̱̇ā̄n**^{ʔ}**V** 'louse' (I am not convinced by deriving this Basque word from a "Romance" source; I would rather suppose the reverse direction of borrowing).

(32) Bq **astigar** 'maple tree': Proto-Nakh (not just Batsbi) ***stagar** 'Acer platanoides'. The word has also a Hurrian parallel: **taškar-innə** 'Buxus sempervirens' (borrowed in Akkad. **taskarinnu**), see Diakonoff—Starostin 1986:25. The presumed loss of initial velar in the Basque word does not convince me (two cases — this and **azal** 'skin', see below — are certainly not enough for postulating a phonetic law, especially since the conditions for such a loss are not at all clear).

(33) Bq **asto** 'donkey': PWC ***č**^{w}**VdV** 'donkey' (Chirikba's etymology). I can certainly not accept the derivation of **asto** < ***hartz-to** 'small bear'. (By the way: are the semantics any better here than the comparison of "bear" and "hamster" by J. Bengtson, of which Trask makes such fun [see p. 38]? Traditional bascological semantics sometimes produces a very strange impression!)

(35) Bq **atso** 'old woman': PNC ***c̱wŏjV** 'woman, female'.

(39) Bq **axeri**, **azari** 'fox', Spanish **zorro**: PNC ***chwōlĕ** 'fox'. In spite of all Michelena's and Trask's arguments, I think it is too difficult to separate the Basque and the Spanish words (maybe the source of the Spanish word was not the Basque form itself, but a form of some unattested related language), and trace the Basque form to a Latin personal name. I can understand, too, that no modern dialects have here any trace of nasalization needed to reconstruct a form like ***azenari**.

(42) Bq **-ba** (in kinship terms): PWC ***pa** 'son, daughter', PY ***pu**^{ʔ}**-** 'son, daughter', PST ***PV-n** 'grandson, nephew'.

(49) Bq **baso** 'woods': PNC ***wɨ̄cV** 'mountain'

(54) Bq **behazun** 'bile': PNC ***c̱wäjmĕ** 'bile', PY ***seŋ**, PST ***sin** 'liver'. [Attempts to derive the Basque word from 'eye', mentioned by Trask, are another example of very curious historical semantics!]

(55) Bq **beko** 'forehead, beak': PEC ***bĕḳwə** 'part of face, mouth'. Here again, we have late Latin BECCU 'beak, bill' of obscure origin — quite probably borrowed itself.

(73) Bq **bizar** 'beard': PEC ***mēć̣uri** 'beard'. Here Trask scolds Bengtson and Chirikba for not having removed the "body part prefix". His main argument is: "You can't have it both ways!" But this is a strange argument: we all know that English has a prefix *be-* (*believe*, *begin*, *bespeak*, *beside*, etc.); shall we insist that the same prefix is present in *beggar*, *beetle*, or *beverage*? [The same argument is given in J. Bengtson's reply to Trask.]

(75) Bq **bizkar** 'back': Abkh. **a-zk**^{w}**a** 'back', PY ***suga** 'back (adv.)'.

(84) Bq **egur** 'firewood': PNC ***gōrV** 'stick, pole'.

(86) Bq **ekei** 'material' (Trask amends it to **gai** ~ **gei**, which really makes no difference for the comparison): PEC ***q̱̇wăjV** 'thing, things'.

(91) Bq **entzun** ‘hear, listen’ (even if reconstructed as ***enezun**): PNC ***=ămc̣E** ‘know’, PY ***ˀVt-** ‘know’, PST ***siə** ‘know, think’.

(92) Bq **erdi** ‘half, middle’: PNC ***=ĕƛ̱’E** ‘middle, half’, PY ***ˀa(ˀ)l** ‘half’, PST ***ƛăj** ‘middle’. This is one of a series of interesting examples in ***-ƛ’-** > **-rd-** by J. Bengtson recently. (See his comments on p. 91 of *Mother Tongue I.*)

(98) Bq **esne** ‘milk’: PNC ***s̱ănˀu**, PY ***de(ˀ)n** (in NCED 982 we have reconstructed ***s̱ămˀV**, but **-m-** is hard to distinguish from **-n-** in this position; see discussion in Starostin 1995:220).

(105) Bq **gal-du** ‘lose’: PNC ***=igwVł** ‘lose’.

(120) Bq **gose** ‘hungry’: PNC ***gašē** ‘hunger’.

(138) Bq **hauzo** ‘neighbor’: PNC ***HV̄č̱wE** ‘guest’ (cf. also PY ***ˀəča** ‘guest’, see Starostin 1995:191).

(141) Bq **hegi** ‘ridge’: PNC ***ɦiwə̄rq̱ē** ‘mountain ridge’.

(142) Bq **herri** ‘inhabited place, people’: PNC ***ʡwə̆hri** ‘people, troop’, PY ***har-** ‘name of a people; slave(s)’, PST ***raH** ‘troop’. ‘People’ and ‘troop, army’ are very close semantically; Trask’s objections are unclear to me.

(150) Bq **hogei** ‘twenty’: PNC ***G̱ə̆**, PY ***ˀeˀk** ‘twenty’ (not ***χɔGa** ‘ten’, as in Ruhlen—Starostin 1994:270).

(154) Bq **hur** ‘hazelnut’: PNC ***ˀwā̆rƛ̱’V** ‘nut’.

(167) Bq **ipini** ‘put’: PNC ***ˀiman-** ‘stay’. Trask for some reason thinks that eliminating ***m** from Proto-Basque destroys all comparisons with roots having ***m** in other languages. Why? — if original ***m** changed to ***p** or ***b** in Proto-Basque, all the comparisons still hold!

(170) Bq **itain** ‘tick’: PNC ***ṭaHnā** ‘nit’. In spite of Michelena and Trask, I really cannot imagine how **itain** can be a result of contamination of **akain** and **izain** (?)

(173) Bq **itsu** ‘blind’ : PEC ***=Vc̱V** ‘dark, blind’ (vowels are hard to reconstruct here, see NCED 1017).

(175) Bq **izan** ‘be’: PNC ***=ă̌sA** ‘sit, stay’ (***=ŏsV** is just an ablaut variant, see NCED 281-282), PY ***hVs-** ‘be’.

(177) Bq **ize-ba** ‘aunt’: PEC ***=īlć̱wi** ‘girl, woman’.

(188) Bq **josi** ‘sew’: PNC ***=irś̱E** ‘weave’ (Chirikba’s etymology).

(197) Bq **karats** ‘bitter’: PNC ***q̇ĕɦlV**, PY ***qV-qVr-**, PST ***ghāH** ‘bitter’.

(198) Bq **kokot** ‘nape’: PY ***kəqənt-**‘neck’; cf. PNC ***GwVnG̱wV** ‘throat’, PST ***Gə̄ŋ** ‘neck’; see Starostin 1995:237. The word of course is expressive, but the Romance forms (Old Spanish **cocote**, etc.) are highly suspicious to me as a probable source of the Basque word (rather the other way round).

(199) Bq **korotz** 'dung': PNC ***ḳurč̣V** 'dung'. Deriving the Basque word from Latin CROCEA 'saffron-colored' (Corominas—Pascual 1980) seems to me perhaps the highest peak of historical semantics in Basque etymology. See also the lengthy discussion in Trask's response (p. 176).

(206) Bq **lau(r)** 'four': PWC ***p-ƛ̄'ə** (Chirikba's etymology). Cf. also PST ***P-lij** 'four'.

(209) Bq **magal** 'lap': PNC ***bVnḳwA** 'belly' (Chirikba's Basque-WC comparison; see NCED 318).

(220) Bq **muin** 'brain, marrow': PEC ***mă̆ƒinɨ̄** 'brain' (for this reconstruction, see NCED 797-798). Trask reconstructs Proto-Basque as ***bune**, which is quite fine for the correspondence (see above, under # 167); however, tracing it back to Latin FUNE 'rope' is again a great achievement of historical semantics.

(228) Bq **nigar** 'tear': PEC ***nĕwq̇ŭ** 'tear, pus' (cf. also PY ***dɔkŋ** 'pus', PST ***nōk** / ***nōŋ** 'pus').

(238) Bq **osin** 'deep place in a river': PY ***sin-** 'well, spring', PEC ***ˀwɨc̱ɨnV** > ***ˀwɨnc̱V** , PST ***CĕŋH** 'well, spring'.

(240) Bq **oso** 'whole, complete': PNC ***ƒiŏc̣V**, PY ***ˀute** 'full, fill'.

(241) Bq **otso** 'wolf' (if < ***Potso**): PNC ***bħĕrc̣ĭ** 'wolf', PY ***pes-tap** 'wolverine', OCh. ***prāts** 'a mythic predator'.

(251) Bq **sabel** 'stomach, abdomen': PNC ***ǯăbV** 'kidney', PY ***tVpVl'-** 'spleen', PST ***č(h)uap** 'lung'. See Starostin 1995:292.

(253) Bq **sagu** 'mouse': PNC ***cārgwɨ̄** 'weasel, marten', PY ***saˀqa** 'squirrel'. This comparison seems impeccable to me, and I do not know what is more impressive in the comparison of Bq **sagu** with Georgian **tagvi** 'mouse' (Bouda, Michelena, Trask).

(261) Bq **sits** 'moth': PEC ***s̱wĕʒV** 'a biting insect' (NCED 988-989). Trask does not know what to do with the Basque homonym **sits** 'dung': I think it is just a fortuitous homonym.

(275) Bq **txori** 'bird': PEC ***č̣HwīlV** 'small bird'. Again, Basque etymology demonstrates wonders here. I can understand (taking into account the Roman tradition) the shift from 'bird' to 'fortune', but hardly vice versa!

(276) Bq **txorru** 'root (of a hair)': PEC ***ć̣ħwərV** 'hair'. While **tx-** may be unetymological (thus the original form may have been something like ***zorru**), it is hard for me to imagine the meaning 'root of a hair' going back to 'sack, bag', as suggested by Trask.

(284) Bq **ur** 'water': PEC ***ħwirɨ** 'river, lake'.

(289) Bq **xahu** 'clean' (< ***sanu**): PNC ***ʒ(w)EnHV**, PY ***tur-**, PST ***siaŋ** (/***ciaŋ**). (See a discussion of these forms in Starostin 1995:290; in NCED 552, the PNC reconstruction is somewhat different: ***Hăʒ́Em-**.) Michelena and Trask derive the Basque form from Latin SANU 'healthy', which is another tour de force of semantic imagination.

(292) Bq **zahar** 'old': PNC ***swĕrho** 'old'.

(293) Bq **zain** ‘nerve, blood vessel’: PNC ***s̱ēħmV** ‘vein, muscle’.

(304) Bq **zigar** ‘mite’: PNC ***ḉäkwə̆** ‘flea, biting insect’.

(306) Bq **zikiro** ‘castrated goat’: PNC ***ʒĭkV** ‘goat, kid’.

(312) Bq **zu** ‘you’: PNC ***źwV** ‘you’.

(315) Bq **zur** ‘wood’: PNC ***c̣wɨ̄łɦV** ‘stick, tree’.

(316) Bq **zuri** ‘white’: PNC ***Hč̣wā̆rV̄** ‘gray, yellow’.

The 54 listed etymologies seem quite satisfactory and interesting to me. Even Trask has no objections in most cases (## 35, 84, 86, 98, 105, 138, 150, 154, 177, 188, 228, 284, 293, 304, 306, 312, 315, 316), and only minor semantic adjustments in many others. How does it correlate with Trask’s final conclusion: “...the evidence so painstakingly assembled for relating Basque to the other ‘Dene-Caucasian’ languages amounts to precisely zero”? What sort of math should one use to get an equation “54 = 0”? And this is still not all — let us move further.

4. Finally, there is the most significant group: a small number of comparisons pertaining to the most basic vocabulary. I shall list all such cases below:

(1) Bq **a** ‘article and demonstrative pronoun’ (the full stem, as Trask states, is **har-**, **hai-**), compared by Chirikba with PWC ***a**. PNC had both ***ha-** and ***ʔŏ** (see NCED 218). Of course, pronouns like this are rather universal, but the comparison still holds. [Let me point out here that the mere fact of a root’s presence in many linguistic families does not mean that the comparison is wrong: this is just the sort of twisted “anti-omnicomparativistic” logic that puts skeptics on a search for some mystic “other explanations” whenever they see a root widely enough spread in different phyla.] Cf. also Bq **hau** etc. in (135).

(4) Bq **adar** ‘horn’: PEC ***ƛ̱wɨ̆rV**. This comparison is interesting, but I am not quite sure about the correspondence of Basque **d** to the EC lateral.

(7) Bq **agor** ‘dry’: PNC ***=iG̱wAr**, PY ***qɔr$_1$-**. The forms **igar**, **eihar** may be unrelated (as Trask says), but **agor** is still there (as he himself states).

(11) Bq **aho** ‘mouth’: PY ***χowe**, PST ***KhōH**. Basque **-h-** may be secondary (as per Jacobsen’s discussion) and the word may be reconstructed in Proto-Basque as ***ago**, but the comparison is still extremely interesting (I compared the PY and PST forms with PEC ***ħḳwĕłV̄**, as per Starostin 1995:303, but I may agree with Blažek that this comparison should be rejected — because of the unclear **-ł-** in PEC).

(38) Bq **azal** ‘skin’: PWC ***c^{w}a** ‘skin’. The PWC form actually goes back to PNC ***ʔwā̆rćwə̆** ‘skin’ (as per NCED 228), which strengthens the case of this etymology (proposed by Chirikba). I am not convinced by Trask’s speculation about the Basque form originally having been ***kazal** (how does it agree with the critic’s own statement that “no native Basque word could begin with any of *p*, *t*, *k*, *d*, or *r*”?).

(57) Bq **belarri** ‘ear’: PNC ***łĕHłe** ‘ear’. (cf. also Hurr. **lēla** ‘ear’, see NCED 756). Although Trask argues hard against the Basque “body part” prefix **b(e)-**, I find the evidence rather convincing.

(83) Bq **egun** 'day, sun': PNC ***Hwīq̇V** 'day' (frequently with a nasal suffix, ***Hwīq̇VnV**), ***wirəq̇A** 'sun' (probably an original compound ***hwēri** 'day' + ***Hwīq̇V** 'day, sun'), PY ***xiˀG** 'day', ***xiGa** 'sun', PST ***x(r)ŏk** 'light, sunrise'; see a discussion in Starostin 1995:296.

(103) Bq **ez** 'not, no': PEC***ʒ́ə̆** 'not, no'. (NCED 1101-1102).

(122) Bq **haragi** 'meat, flesh': PNC ***räƛ̱'ɨ̆** 'meat, flesh'.

(145) Bq **h-**, **-k** (***-ga**) 'thou': PNC***ʁwV̄** 'thou', PY ***kV-**, ***ˀVk-** 'thou'.

(146) Bq **hil** 'dead, die, kill': PNC ***=iwƛ'E** 'die, kill'.

(176) Bq **izar** 'star': PNC ***ʒwhărī** 'star' (cf. also PST ***sēŋ** 'star')

(178) Bq **izen** 'name': PNC ***ʒwĕrhɨ̆** 'name' (cf. also OCh. ***seŋ** 'name')

(185) Bq **jakin** 'know': PNC ***=ɨ̆q̇Ē** 'know' (cf. PST ***qēn** 'to see, know')

(187) Bq **joan** 'go': PEC ***=VˀwVn** 'go', PY ***hejVŋ** 'go', PST ***ˀwă(ŋ)** 'go'

(227) Bq **ni** 'I': PEC ***nɨ̆** (cf. also PY ***-ŋ** , PST ***ŋa** 'I')

(229) Bq **no-** 'interrogative stem, who': PEC ***hī-nV-** 'interrogative stem', PY ***ˀan-** 'interrogative stem', PST ***nāiŋ** 'interrogative stem'

(236) Bq **or** 'dog': PNC ***χHwĕje**, obl. stem ***χHwĕjrV-**, see NCED 1074; PST ***qhʷīj** 'dog'. It should be noted that if this comparison is correct, then Basque reflects an original oblique case form (just like some EC languages: Khinalug **pχra** , Kryz **χʷar**, etc.).

(263) Bq **su** 'fire': PNC ***c̣ăjɨ̆** 'fire'

(302) Bq **ze-** 'interrogative stem, what': PNC ***s̱aj** 'what', PY ***ˀas-** /***sV-** 'what'; cf. also PST ***su** 'who'.

Well, it should be clear by now that I picked out of Bengtson's data words belonging to Swadesh's 100-word list. Amazingly enough, there are almost twenty of them (and I have omitted several cases like Bq **mihi** 'tongue' and **begi** 'eye' where the phonological side is not clear to me), with exact semantic matching between Basque and North Caucasian. I am not going to discuss glottochronology right now, but any reasonable statistician would of course say that this is a significant result. The same would be clear for any unbiased comparative linguist. The latter would, of course, also demand a set of regular correspondences — which I think will be also ultimately possible, after we collect all the existing data, throwing out faulty comparisons and taking into account all which is known about Basque historical phonetics. The uncertainty of comparative phonology is the only factor that still keeps me from enthusiastically including Basque in Sino-Caucasian (or Dene-Caucasian).

Now, in his final response (p. 196-197), R. L. Trask decided to act as an "ambitious long-ranger" and produced a list of Basque-English comparisons — to convince everyone that Basque can be compared to any other language with equal success. Let us inspect it more closely.

Out of Swadesh's 100-wordlist, we have three exact semantic matches (assuming everything is OK with phonology): Bq **bi-hotz** 'heart' ~ Eng **heart**, Bq **sa-bel** 'belly' ~ Eng **belly**, and Bq **bel-tz** 'black' ~ Eng **black**. To be sure, there are also other 100-wordlist items, but only with different meanings (Bq **bide** 'road' ~ Eng **path**, Bq **buru** 'head' ~ Eng **brow**, Bq **gibel** 'liver' ~ Eng **giblet**, Bq **j-oan** 'go' ~ Eng **wend**). I am sure that Trask would have included all basic look-alikes with identical meaning into his "hurried list of comparison". But there are just three, well below the threshold of accidental correspondences.

What is even more astonishing is that a significant part of Trask's Basque-English look-alikes are actual cognates! The case of English **muzzle** and Basque **musu** was discussed by Trask himself on p. 54 (according to Trask, both go back to an obscure "Romance" source). There are also Bq **pinu** 'pine', **tratu** 'deal, pact' (Eng **treaty**), **zilar** 'silver' (discussed by Trask on p. 71), **zinu** 'sign', **ziur** 'sure', **zola** 'sole' etc. These are of course borrowed both by Basque and by English from a common source. It seems that after all it is not so easy to find a significant amount of accidental look-alikes in any two languages.

This all should be very clear to any person who studied elementary statistics and tried to apply it to linguistic data. Instead, we have to engage ourselves constantly in repeating the same things all over again and again.

My conclusions are the following (and I sincerely think there can be no others):

1. Basque is quite possibly a member of the Sino-Caucasian (Dene-Caucasian) family.

2. We do not as yet possess enough knowledge about phonological correspondences between Basque and other Sino-Caucasian languages, although recently J. Bengtson undertook some serious steps in this direction. (Even one of the opponents, W. Jacobsen, had to acknowledge this while discussing the **z-** and **s-** distribution on page 139.)

3. I am rather pessimistic about the future of Basque studies. As a matter of fact, I am very worried about the fate of long range comparison as such. Our field is right now very clearly broken in two opposite camps that cannot find common language. On one side, there is a small group of long rangers ("lumpers"), undertaking bold, but very often inaccurate, comparisons. On the other side there is the vast majority of narrow specialists ("splitters") who are for the most part incapable of stepping beyond their own fiefdoms and who spend a lot of time and effort in trying to dismantle — against the obvious — any attempt at establishing deeper genetic links. Imagine how we would all profit from normal cooperation rather than from constant quarrels!

4. Trask quotes the old adage of historical linguistics: "Look for Latin etymologies on the Tiber". Well, if we stuck to that golden rule, we wouldn't know that Latin was related to Russian or Hindi — after all, their homelands are quite far from the Tiber.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Response to Starostin

R. L. Trask
School of Cognitive and Computing Sciences
University of Sussex
Brighton BN1 9QH UK
larryt@cogs.susx.ac.uk

1. Introduction. The editors have kindly invited me to reply to Sergei Starostin's comments on the Basque-Dene-Caucasian debate. I do not find this entirely easy, since Starostin has in large measure chosen merely to reiterate the comparisons and the arguments advanced earlier by others and already discussed by me. Hence there is a danger that I will wind up merely repeating what I have said earlier, an outcome which I believe no one will find interesting or helpful. Therefore, I propose to do the following here. First, I shall respond to the genuinely new points made by Starostin, and to those instances in which I believe he has misconstrued my position. Second, I shall examine his conclusions and try to identify as explicitly as possible just where our differences lie. Third, I shall raise a point which has not so far figured prominently in the debate, but which is, in the eyes of us Vasconists, of critical importance, and substantially destructive of the case which has been made for relating Basque to the other languages cited. Finally, I shall close with a statement of the proper relation between specialist work and comparative investigations by non-specialists.

2. Starostin's comments. Starostin begins with some general remarks, one or two of which I shall reserve for discussion later in this response. Interestingly, he advances a number of criticisms of the comparisons put forward not only by Bengtson, Ruhlen, and Chirikba, but also by earlier workers in a similar vein, including Lafon and Bouda. Specifically, he draws attention to a number of further errors of the type which will be all too familiar to readers who have been following this discussion: nonexistent words cited, words cited wrongly, glossed wrongly, or segmented wrongly, and so on. Unlike me, though, he appears to believe that such grave errors are no more than a minor irritant, and in no way an indictment of the procedure of trying to draw comparisons between languages one is ignorant of. My own view is that it's hard enough to do comparative work on languages you know intimately, and that trying to work on languages you don't know is a recipe for disaster.

Starostin asks why I consider it useful to point out that some of the more recent comparisons are repetitions of those drawn earlier by Lafon, Bouda, and others. My reason is that these comparisons form part of that body of work which was resoundingly dismissed as without value by Michelena a generation ago — and which is now dismissed by Starostin as well on the ground that it was based on "simply abominable" Caucasian materials. If these comparisons were no good in 1950, then they are no better today.

Starostin declines to accept the Romance origin of a number of the Basque words cited as comparanda, preferring to see instead loans into Romance from Basque or from an unattested relative of Basque. I find his position astounding. Let's consider these cases.

(14) Basque *akain* 'tick' and (170) *itain* 'tick'. All specialists seem satisfied that these words are of Romance origin; see the lengthy discussion in Corominas and Pascual (1980). Starostin simply declares that he would "rather suppose" the reverse direction of borrowing. This is not an argument that can be taken seriously: he can suppose whatever he likes, but no one who is familiar with the data agrees with him.

(39) Basque *azeri* 'fox'. Starostin rejects the derivation of this from Latin ASENARIU, on the mysterious ground that he finds it "too difficult" to separate this word from Castilian *zorro* 'fox'. Apparently he has not read what I've written: the evidence for the Latin origin of *azeri* is overwhelming, and this etymology is accepted by everybody who has seen the data. A connection with *zorro* is indefensible: no such word is attested in Basque; the Spanish word is first attested far from the Basque Country, near the Portuguese border; and it is first attested with a *voiced* fricative — hardly possible in a loan from Basque. Starostin wants to reject an utterly well-substantiated etymology in favor of a flight of pure fancy backed up by no evidence at all. For him, it seems, vague resemblances are paramount, and hard evidence is of no consequence.

(198) Basque *kokot* 'nape'. Unbelievably, Starostin wants to see the Romance cognates as borrowed from Basque. But *kokot* could hardly look less like a native Basque word: it violates almost every available phonotactic and morpheme-structure constraint applying to native Basque words (see Trask 1996: sections 3.17-3.18), and there is no known case of a Basque word being borrowed so widely into Romance. Starostin's position is nothing but fantasy, and he will find no Vasconist or Romanist anywhere who would take his idea seriously for a moment.

(199) Basque *korotz* 'dung'. This word too looks like anything but a native Basque word. Starostin expresses amazement that this word could be derived from Latin CROCEA 'saffron-colored', but this is only the *ultimate* source proposed by Corominas. The *direct* source is the Old Castilian reflex *croça,* which, among other things, means 'dung covering on a haystack'. And where is the semantic problem in this?

(220) Basque *muin* (and variants) 'pith, marrow, inner part'. The numerous variants point to an original **bune* or **fune,* and Latin FUNE 'rope' is a phonologically perfect source, even if the semantics is admittedly difficult. (But note that the Basque word is also attested as 'spinal cord'.) In any case, *muin* does not mean 'brain': 'brain' is *burumuin* 'head-interior', 'head-pith', from *buru* 'head'; in context, this is occasionally reduced to *muin,* just as English *eyelashes* and *fingernail* are sometimes reduced to *lashes* and *nail.* It is completely out of order to cite the Basque word as 'brain' purely in order to get a match which doesn't otherwise exist.

(239) Basque *xahu* 'clean'. Starostin describes the accepted etymology of this, from Latin SANU 'healthy', as "another tour de force of semantic imagination". For heaven's sake: where does he think English *sanitary* 'clean', *sanitize* 'make clean', and *sanitation* 'keeping things clean' come from? Martian?

There are a very few other comments by Starostin which merit a response.

(33) Basque *asto* 'donkey'. Starostin queries Azkue's proposed derivation from *hartz* 'bear' + *-to* diminutive. True, the semantics is difficult, and I wouldn't want to declare this etymology obviously right. But the phonology is perfect, so we can't rule it out, either. Anyway, the existence of the Roncalese variant *arsto,* completely understandable in Azkue's terms, is highly embarrassing for Starostin's West Caucasian comparison.

(167) Basque *ipini* 'put'. Starostin declares "Trask for some reason thinks that eliminating **m* from Proto-Basque destroys all comparisons with roots having **m* in other languages." Not so. What I have objected to, very strongly, is the insistence of Bengtson and others on pretending that an */m/ was present in Pre-Basque in order to make dubious comparisons look better. Starostin goes on to declare expansively that "if original **m* changed to **p* or **b* in Proto-Basque, all the comparisons still hold!" [his punctuation]. Yes, and if any number of other hypothetical but convenient changes had occurred in some ancestral form of Basque, all of the comparisons would look much better than they do. I do not believe it is possible to do comparative linguistics by inventing hypothetical and purely *ad hoc* developments in unattested languages in order to get the result required by some hopeful comparativist.

(275) Basque *txori* 'bird'. Starostin's remarks here are incomprehensible. The Basque word is a diminutive of original **zori,* which itself survives as the word which earlier meant 'omen' but now means 'luck, fortune'. Starostin says the following. "Again, Basque etymology demonstrates wonders here. I can understand...the shift from 'bird' to 'fortune', but hardly vice versa!" [his punctuation]. But no one has suggested vice versa: what he can understand is precisely what we are asserting, so where is the problem?

In sum, then, Starostin is rejecting well-documented and universally accepted etymologies in favor of his own intuitions and empty speculations. And why should anyone take this seriously?

A further curious feature of Starostin's commentary is his appeal to statistics. In one place, he dismisses the exacting statistical work of Don Ringe as "strange" — but without making the slightest attempt at countering Ringe's case. In another place, however, he makes the following declaration: "Statistical considerations tell us that to be considered related languages must have no less than 10% of their most basic vocabulary in common (within Swadesh's 100-wordlist)." Now, where does he get this novel principle from, and just what "statistical considerations" underpin it?

Lexicostatistics was not designed to identify previously undetected genetic links, and it cannot be properly used for that purpose. The technique was invented in order to provide a measure of linguistic distance between languages which are *already* known to be related and between which secure cognates have *already* been identified. But Starostin here is trying to do something very different, and totally inappropriate. What he's doing is working through the Swadesh word-list for some languages of his own choosing, languages not known to be related at all, and looking for miscellaneous resemblances. If he finds some (as of course he does), he declares the result to be interesting; by his own arbitrary criterion, if he finds ten miscellaneous resemblances (out of a hundred) which strike *him personally* as impressive, he announces that he has found evidence of a genetic link.

Sorry, but this is nonsense. In the present case, for example, he declares a match between Basque *or* 'dog' and PNC **χHwĕje 'dog'* (oblique stem **χHwĕjrV-),* one between Basque *egun* 'day' and PNC **Hwīq̱̇qV* 'day', one between Basque *azal* 'skin' and PNC **ʔwā̆rćwə̆* 'skin', one between Basque *belarri* 'ear' (probably from **berarri)* and PNC **łĕHłe* 'ear', and so on and so on. He then announces delightedly that he has found "almost twenty" matches between Basque and PNC, and that this is more than enough to satisfy his criterion.

But look. Suppose I deny that these are persuasive matches (and I certainly do). How can Starostin respond? As far as I can see, all he can do is to declare grandly that *he personally* is impressed by them, even if I'm not. He has *no criterion* to appeal to in order to support his case, no way of conducting a rational discourse: for Starostin, these words are impressively similar because he says they are, and that's the end of it. But is this really a scientific way of proceeding? Is this his idea of rigorous statistical analysis, vastly superior to the "strange" approach of Ringe? I don't think many readers will be impressed by a methodology which depends on the personal opinions and intuitions of one man and on nothing else whatever.

Starostin goes on to complain that my joke Basque-English comparison is a failure because it doesn't reach his criterion of ten matches in the Swadesh 100-word list. True, it doesn't, because (among other things) I hadn't realized that these were the rules we were playing by. Certainly Bengtson, Ruhlen, and Chirikba never appealed to any such criterion, and in fact they pay no attention to the Swadesh word list. OK; then; let's play by Starostin's rules. Here are some Basque-English comparisons from the Swadesh 100-word list.

Basque	English
hauts 'ashes'	*ash*
sabel 'belly'	*belly*
beltz 'black'	*black*
odol 'blood'	*blood*
hotz 'cold'	*cold*
idor 'dry'	*dry*
bete 'full'	*full*
(bi)hotz 'heart'	*heart*
hil 'kill'	*kill*
(j)akin 'know'	*know*
luze 'long'	*long*
mendi 'mountain'	*mount(ain)*
euri 'rain'	*rain*
erro 'root'	*root*
hondar 'sand'	*sand*
(e)san 'say'	*say*
(e)seri 'sit'	*sit*
izar 'star'	*star*
bero 'hot'	*warm*
gu 'we'	*we*

Following Starostin's lead, I will now declare that I find these matches persuasive. Since I have twenty persuasive matches out of a hundred words, then, by Starostin's criterion, I have an excellent case that Basque is genetically related to English. Of course, he may not agree with my judgments of similarity, but then I don't agree with his, either — but what can we possibly do about it in either case? If he can base his case entirely on his judgments, then I can base my case on my judgments, and there is no basis for discussion. So much for Starostin's case.

(Oh, it's true, of course, that English *mountain* is a loan word, but then so are many of the Basque words solemnly adduced in comparisons by Starostin, Bengtson, and others.)

3. Starostin's conclusions. Starostin offers four conclusions, adding that he "sincerely think[s] there can be no others".

1. "Basque is quite possibly a member of the Sino-Caucasian...family."

   What sort of conclusion is this? Any language is quite possibly related to any other language, and indeed all languages are quite possibly very remotely related. The only interesting question is whether we have any persuasive evidence for accepting a relationship in a particular case. In the case of Basque and Dene-Caucasian, my answer is an unambiguous "no".

2. "We do not as yet possess enough knowledge about phonological correspondences between Basque and other Sino-Caucasian languages."

   What phonological correspondences? I would put this more starkly: we do not possess any phonological correspondences between Basque and any of the other languages cited.

3. "I am rather pessimistic about the future of Basque studies."

   Eh? Forgive my bluntness, but does Starostin actually *know* anything about the state of Basque studies? I would maintain that Basque historical studies are currently going through one of the most vigorous and rewarding periods in their history. In the last few years, we have made enormous progress in understanding the prehistory of nominal and (more especially) verbal morphology — for example, we now understand a great deal about the origin of those mysterious "third-person" prefixes in the verb; we understand how the *e*-class of verbs came into being, with all its idiosyncrasies; we know important things about the origin of the local cases; we have largely worked out the historical development of the various non-finite verb-forms; and we have worked out the approximate order in which the various bits of morphology were incorporated into the finite verb. We have obtained a significant number of new etymologies. We have worked out the major morpheme-structure constraints applying in Pre-Basque. And these few examples are far from being the whole story. So why is Starostin pessimistic about our efforts? Merely because we decline to accept hopeful comparisons that fly in the face of all our results?

4. If we took the old adage "Look for Latin etymologies on the Tiber" seriously, "we wouldn't know that Latin was related to Russian or Hindi."

   Nonsense. Starostin must surely realize that this is a gross perversion of the sense of the adage. This adage does not mean "No language ever has distant relatives or distant cognates"; that's absurd. It means merely that, when seeking the source of a particular word, we should first consider the most obvious sources — language-internal developments and loans from neighboring languages — before we get excited about vaguely similar-looking words in languages which lie thousands of miles away and are not known to be related to it. The adage has proved its value in over 200 years of scholarly work, but Starostin, apparently, would have us scrap it in favor of a new dictum: look for Latin etymologies in the Himalayas (or perhaps better the Caucasus). That this is so is shown indisputably by his practice of rejecting even the most blatant and secure Latin and Romance etymologies for Basque words in favor of fanciful connections with languages spoken many thousands of miles away.

4. The evaluation of comparisons. Starostin here continues the practice of other long-rangers in assuming that demonstrating a genetic link between languages consists merely of finding lists of similar-looking words and morphs. But, as I have argued at length, such lists can always be found between any arbitrary languages, and they mean nothing unless backed up by at least one further prop: either systematic phonological correspondences or reconstructible grammatical systems. In the case of Basque and Dene-Caucasian, we certainly have no systematic correspondences. So, then: do we have any reconstructible grammar? Is there any evidence that we can reconstruct a common grammatical system for the putative common ancestor of this alleged family, with identifiable remnants of that ancestral grammar surviving in the various daughters?

As far as Basque is concerned, the answer to this question is plainly "no". The sole effort known to me along these lines is Bengtson's attempt to interpret the frequent initial vowels of Basque as fossilized cognates of the noun-class prefixes of North Caucasian languages — but he hasn't been able to make this work. If the proponents of Dene-Caucasian have made any progress in reconstructing a grammatical system for Proto-Dene-Caucasian, apparently Basque doesn't fit.

Very well, then. Have the Dene-Caucasian comparisons shed *any light at all* upon the outstanding problems in Basque prehistory? None. Zero. Nothing at all. Not one single point of Basque phonology or grammar receives even the faintest illumination from the Dene-Caucasian comparisons on offer. All we ever get is the observation that Basque has some words that look a bit like some words in the other languages, which is neither surprising nor interesting.

But there are any number of puzzles which we cannot solve with the Basque data alone, and it is here that we Vasconists would expect to find some assistance offered by any valid genetic proposal. If these other languages are truly related to Basque at any discoverable level, then we would expect to find that they share with Basque identifiable remnants of an ancestral phonology and grammar, remnants which would shed some light on these puzzles. But this never happens: so far as we are concerned, the Dene-Caucasian comparisons have been unrelievedly sterile.

Here is a brief list of some of these puzzles.

1. Pre-Basque clearly had an extraordinarily large proportion of lexical items beginning with a vowel, and it permitted only a very few word-initial consonants. Why is this so? Did some ancestral form of the language undergo a systematic loss of certain initial consonants? Do any other languages exhibit something similar, or do they exhibit any initial consonants whose loss might have produced what we see in Pre-Basque?

2. The oldest stratum of finite verbal morphology that we can recover is absolutive number marking, which is older even than person agreement. Absolutive plural marking is pervasive, occurring in all verbs and in all three persons. It is also very messy: a number of quite different morphs are used to mark plurality, and these sometimes follow the verbal root, sometimes precede it, and sometimes appear to be infixed into it. Examples: *da* 's/he is', *dira* 'they are'; *dago* 's/he is', *daude* 'they are'; *datza* 's/he lies down', *dautza* 'they lie down'; *doa* 's/he goes', *doaz* 'they go'; *dabil* 's/he is busy', *dabiltza* 'they are busy'; *du* 's/he has it', *ditu* 's/he has them'; *dauka* 's/he has it', *dauzka* 's/he has them'. Indeed, so messy is this system that I am inclined to wonder whether some remote ancestor might have had stem-suppletion for absolutive number. And it is certainly very old. Do any other languages display anything remotely similar?

3. A finite verb has a prefix slot which must always be filled. Whenever a suitable first- or second-person argument is present in the sentence, this prefix slot is filled by an agreement marker which is plainly cognate with the corresponding free pronoun and probably derives from a cliticized pronoun. But, whenever no suitable NP is present to trigger this agreement, the prefix slot is filled instead by a marker of tense or mood: *d-* in the present, *z-* in the past (Bizkaian zero, which is probably older), *l-* in the irrealis, and *b-* in the jussive. These prefixes appear to be the second-oldest feature of finite verbal morphology. Do any other languages show anything like this system, or at least plausible etyma for these prefixes?

4. In the first- and second-person plural, both the agreement prefixes and the agreement suffixes are clearly cognate with the corresponding free pronouns: *gu* 'we', agreement markers *g-* and *-gu*; *zu* 'you', markers *z-* and *-zu*. Surely we are looking here at cliticized pronouns. But the singular is much more complicated. Only the agreement prefixes are clearly cognate with the pronouns, while the suffixes are different: for example, *ni* 'I', prefix *n-* but suffix **-da*. The second person is even worse, since the suffix varies to mark the sex of the addressee: *hi* 'you', prefix *h-*, but suffixes **-ga* (male) and **-na* (female) (or possibly

**-naga)*. These suffixes are inexplicable, and the sex-marking in particular is absolutely isolated and enigmatic in Basque. Do the singular forms perhaps retain traces of ancient stem-alternations in the singular pronouns? Do any other languages exhibit anything similar?

5. Both finite and non-finite verb-forms exhibit a pair of affixes whose original function appears to have been to increase the valency of a verb by adding a dative argument; these are *-ki* ~ *-i* (probably both from **-gi*) and *-ts-*. There is evidence that the first was anciently used with intransitive verbs and the second with transitive verbs, though this is not certain. Strangely, in a language which is exclusively suffixing, postpositional, and SOV, these morphs, in finite forms, precede their associated agreement markers. Where did these things come from? Did they originate as serial verbs, or as adpositions? If so, what are they doing preceding their apparent objects? Do these markers retain evidence of an ancient VO order? Do any other languages have anything like this?

6. Word-forming prefixes are absolutely unattested in Basque. Yet all non-finite forms of ancient indigenous verbs bear a prefix **e-*, which surfaces variously today as *e-*, *i-*, *j-*, or zero: *ebaki* 'cut', *egin* 'make', *ikusi* 'see', *ibili* 'be active', *jan* 'eat', *joan* 'go', *utzi* 'leave' (< *eutzi,* attested). What was the function of this prefix, and how did the language come to be using a prefix at all when prefixes are otherwise absent? Do any other languages show anything similar?

7. The modern genitive case-suffix is *-en*. But the personal pronouns show *-e* instead, and a number of fossilized forms involving demonstratives also exhibit this *-e*. Is *-en* therefore an innovation replacing original *-e*? If so, where did *-en* come from? Do any other languages show anything similar?

These points constitute no more than a reasonable sample of the puzzles we appear to be unable to solve without more data than we have available. In spite of the repeated suggestions by injured long-rangers that we Vasconists are hostile on principle to proposed genetic links, the truth is that we would fall with tears of joy on any proposal which succeeded in shedding some genuine light on some of these puzzles. So far we haven't seen any such proposals, and the Basque-Dene-Caucasian hypothesis is no more illuminating than any of its numerous rivals. If Starostin or anyone else can show that the North Caucasian languages, or any other languages, share with Basque clear remnants of a single ancestral linguistic system, we'll be buying the drinks. But we will never accept any proposal which simply rejects our most substantial and secure results in favor of a handful of vague resemblances between words.

5. The relation between specialist work and comparisons by nonspecialists. In his conclusions, Starostin, echoing earlier remarks by Bengtson and other long-rangers, takes me to task for my allegedly negative and obstructive attitude toward long-range comparison, and pleads with me to engage in "normal cooperation". This point deserves some serious consideration.

In any discussions between specialists and comparativists, it is perfectly clear what the role of the specialists should be: to provide the most comprehensive and reliable information possible about the prehistory of their specialist languages or families. No responsible comparativist would want to work with scanty, defective, outdated, or unreliable materials if better information were readily available. Starostin himself implicitly recognizes this in his comments. He points out that Illich-Svitych's Nostratic work enjoyed the great advantage of being based on reconstructions painstakingly assembled by specialists in most of the families he was working on, and he complains that his own work on Sino-Caucasian has been severely hampered by the lack of comparable reconstructions for any of the families involved. This is all entirely reasonable.

In the case of Basque, more than a century of equally painstaking specialist work has succeeded in uncovering a great deal of information about the prehistory of the language. We now have a comprehensive and meticulously documented reconstruction of the phonology of the Pre-Basque of some 2000 years ago. We have largely worked out the rather severe phonotactic and morpheme-structure constraints which characterized Pre-Basque words. We understand all of the major phonological changes which have affected the language since the Pre-Basque period. In most cases we can readily distinguish loan words from native words, and we have identified good Latin and Romance sources for the majority of the loan words. We have established that Basque loans into Romance are rare and entirely confined to those Romance varieties immediately adjacent to the Basque-speaking region. We know a great deal about the history of word-formation in Basque, and we can confidently identify purely Basque-internal sources for a large number of compound and derived words.

Therefore we might reasonably expect that any long-rangers eager to locate distant relatives for Basque would fall joyfully upon our results and make heavy use of them in their comparisons, just as Illich-Svitych did with the results at his disposal. But do they do this? Do they, hell.

In fact, they do precisely the opposite. Having already decided, on the basis of hopelessly inadequate and error-ridden information, that they have found some relatives for Basque, they merely wave away all our results as inconvenient. We tell them that native Basque words do not begin with any of /p t k d/; they deny it, and cite modern words with these initials freely in their comparisons. We tell them that native Basque words do not contain /m/, except in limited and identifiable circumstances; they deny it, and cite modern words containing /m/ with equal freedom. We tell them that certain modern words had a very different shape in Pre-Basque; they deny it, and use the modern forms exclusively in their comparisons, because these provide matches that otherwise wouldn't work. We tell them that certain words are transparent loan words, often with readily identifiable Romance sources; they deny it, and insist that the Romance cognates must be borrowed from Basque. We tell them that certain words are obviously of imitative, expressive or nursery origin and hence not available for comparison; they deny it, and cite such words with complete freedom. We tell them that certain words are transparent compounds or derivatives formed within Basque; they deny it, and treat these words as ancient and monomorphemic in order to get matches.

When we protest at all this, they accuse us indignantly of being negative and obstructive, of clinging mindlessly to unsubstantiated dogmas, of having a mind-set that refuses to acknowledge the existence of any relatives for Basque as a matter of principle. Finally, on top of all this, they accuse us of a lack of cooperation.

Readers may decide for themselves just who is guilty of a lack of cooperation in this case.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Response to Trask

Sergei A. Starostin
Russian State University of the Humanities
University of Michigan, Ann Arbor

starosti@umichedu / sergei@starostin.msksu

This discussion seems to go on and on, but in Trask's response, I see some points that I would like to dwell on. I shall quote parts of his reply, and give my comments.

1.

> ... he appears to believe that such grave errors are no more than a minor irritant, and in no way an indictment of the procedure of trying to draw comparisons between languages one is ignorant of. My own view is that it's hard enough to do comparative work on languages you know intimately, and that trying to work on languages you don't know is a recipe for disaster.

I gave this long quotation because I think it is very important in revealing the positions where we all stand. My question to Trask would be: how many languages can one know intimately? Even the best polyglots, to my knowledge, rarely know more than a hundred languages, and of course not all of them "intimately". I have rarely met a historical linguist intimately acquainted with more than a dozen languages.

Now, according to different evaluations, there are from four to six thousand languages on our planet. The immediate conclusion from Trask's statement is exactly what he is talking about throughout the whole discussion: no long-range comparison is possible just because no person can be acquainted with that many languages. This concept actually lies behind all the tide of skepticism towards Nostratic, Sino-Caucasian, all other attempts of long-range comparison: how can you do that if you do not know the languages?

Well, some of us indeed do know the languages, but this is not even important. A requirement that any scholar involved in long-range comparison should possess intimate knowledge of at least a hundred languages seems absurd to me, and adopting such a requirement would necessarily result in complete oblivion of this field of studies. In a similar way, we would require abandoning all studies in biological taxonomy just because no single person can be a specialist both in the taxonomy of mammals and in the taxonomy of invertebrates.

What should be required, of course, is an ability to deal in a competent way with the results of any particular field (in the case of Basque this requirement was not entirely fulfilled by Bengtson, and this is of course his weak point). What is even more important: the results within every particular field (be it Basque, Indo-European, Caucasian, or Khoisan) should be presented by specialists in a way allowing specialists from other fields to use them without having to dig independently (and often incompetently) into those areas. In other words, cooperation should always be kept in mind. And if we start with statements like "trying to work on languages you don't know is a recipe for disaster", then of course no cooperation whatsoever is possible.

2. Concerning the quotations from Lafon, Bouda, and others, I still do not understand it. Let us take an example. Parallel 263 analyzed by Trask is Bq *su* 'fire, hearth': PNC **c̣ajɨ* 'fire'. Again, I quote Trask verbatim:

> Basque *su* means only 'fire', never 'hearth'. The Basque-Caucasian resemblance has been pointed out several times before, for example by Uhlenbeck (1924:23) and by Bouda (1948:10).

I agree that it is not accurate to say that a word means 'hearth' when it doesn't. But 'fire' : 'fire' is still a quite exact semantic match, isn't it? So in this case, there is not the slightest reason to reject the comparison, and the fact that Uhlenbeck and Bouda also noticed it cannot possibly speak against it. The only reason to object to it would be a firm belief that Basque cannot be related to any other language, and, therefore, any similarity must simply be due to chance, even a similarity that cannot be proven by any means to be coincidental.

Let me state my position explicitly once more. I am still not completely persuaded by Basque-Caucasian comparisons — just because I have not yet seen a complete system of phonetic correspondences between Basque and North Caucasian (although some recent attempts by Bengtson seem quite encouraging, and Trask is certainly wrong in saying that "we do not possess any phonological correspondences between Basque and any of the other languages cited", just as he is wrong when he says that "all evidence amounts to precisely zero"; he seems to like these absolute statements). It is theoretically possible that when such a system is presented, it will turn out that we shall have to reject some (or many) of Bengtson's parallels on phonetic grounds, just as it is possible to do in many better known linguistic families. But "destroying" (this is a word that Trask likes) evidence even before it can be evaluated does not seem a proper thing to do.

At this point, all I can do is try to evaluate evidence presented in favor of the Basque-Caucasian relationship and the evidence against it, and try to see which hypothesis works better. It cannot be anything more than a hypothesis before we have the standard tool: phonological correspondences.

3. I think that time will resolve our controversies concerning the etymologies of *akain / itain, azeri, kokot, korotz, muin, xahu*, and *asto* (I am still not convinced by Trask's arguments in those cases). In the cases of *ipini* and *txori*, however, I think Trask just did not see my points.

Trask's assertion is that Proto-Basque did not have any **m*. Now does this mean that Basque cannot be compared with any other language just because all other languages have *m*? Surely, the **m* must have existed and was lost at some point. Now what we actually observe in Bengtson's data is a *regular correspondence* of Caucasian **m* to Basque **p* or **b* (although one must still find out what conditioned this split). I am not saying that this correspondence is indisputable: I just do not see any reason to reject any comparison just because Basque had no *m*. Slavic languages do not have any voiced aspirated consonants: would this be a reason to reject the comparison of Slavic **by-ti* and Old Indian *bhū-*? This seems to me a misunderstanding of some sort.

The case of *txori*. The word means 'bird' in Basque, and 'small bird' in Caucasian languages. Now let me quote Trask once again:

> The Basque word is a diminutive of original **zori,* which itself survives as the word which earlier meant 'omen', but now means 'luck, fortune'.

Doesn't this quotation imply that the word once meant 'omen', then became palatalized (for whatever reason) and this palatalized form meaning presumably 'little omen' changed its meaning to 'bird'? This is exactly what I meant when I wrote that I do understand the shift from 'bird' to 'fortune' (as in Old Spanish), but do not understand the shift from 'fortune' to 'bird'. Despite the advice to look for "Latin etymologies on the Tiber", I shall always prefer a simple comparison ('bird': 'bird') to a sophisticated one unless of course it is very well grounded, which I think is not the case here (this whole story of diminutives in *tx-* seems rather strange to me, especially because you have to accept statements like: "*txorru* 'root [of a hair]' is a diminutive of *zorro* 'sack, bag, pod'" [??]).

4. Statistics. There were a lot of exercises with statistics in recent literature, and not all of them are equally successful. I do, indeed, dismiss the "exacting statistical work" of Don Ringe: a very good criticism of it by W. Baxter and A. Manaster-Ramer has just been published in *Diachronica* (XIII/2:371-384). The authors kindly let me read the manuscript before it was published, and I find almost nothing to add to that on my own.

I am, however, quite prepared to argue for some statistical considerations of my own. The principle that "to be considered related languages must have no less than 10% common most basic vocabulary" is purely empirical: I have done lexicostatistical research on a lot of linguistic families of Eurasia, and I can assure everybody that *all* languages that are usually considered related have more than 10% of cognates within the Swadesh wordlist. It is below that figure where the controversies start (Nostratic, Sino-Caucasian, Austro-Thai, etc., etc.).

I quite agree with Trask that lexicostatistics was not designed to identify previously undetected genetic links; I also agree that it cannot be properly used for that purpose. Trask, however, is completely misjudging my methods here. I have to quote him again:

> What he [Starostin] is doing is working through the Swadesh word-list for some languages of his own choosing, languages not known to be related at all, and looking for miscellaneous resemblances. If he finds some (as of course he does), he declares the result to be interesting; by his own arbitrary criterion, if he finds ten miscellaneous resemblances (out of a hundred) which strike him *personally* as impressive, he announces that he has found evidence of a genetic link.

I am sorry, but this is all wrong. In this whole dispute, I act only as an outsider: I did not choose the languages (Basque and Caucasian), and I am not looking for any resemblances at all. I take all the data from Bengtson's comparisons and do not change anything there. I have never found a single Basque-Caucasian parallel in my whole life. All I am doing is taking Bengtson's comparisons and checking if they are on the 100 word-list or not. So from a scientific point of view, I am trying to keep the experiment as clean as possible — which is, of course, a necessary requirement for all statistical evaluations. I think that this is exactly the "scientific way of proceeding" (to quote Trask once again). Finally, I do not announce that I found evidence of genetic link: I just say that the results are significant, and it is worth while looking for a real proof — which is of course establishing regular correspondences and checking everything once more on a more solid basis.

Now Trask agreed to play by my rules and produced one more list of Basque-English comparisons. All right, let us evaluate this one. To do that, I shall use the rigorous statistical procedure developed by S. Yakhontov. In case our readers are not acquainted with it (it was published only in Russian, and few Westerners read Russian, of course), I will briefly explain Yakhontov's method. Swadesh's 100 word-list is divided by Yakhontov into two parts: a 35-word-list of the most stable words (which are: 'blood, bone, die, dog, ear, egg, eye, fire, fish, full, give, hand, horn, I, know, louse, moon, name, new, nose, one, salt, stone, sun, tail, this, thou, tongue, tooth, two, wind, water, what, who, year') and a 65-word-list of less stable words. Note that Yakhontov uses a somewhat modified 100-word-list (he excluded the words 'all, swim, claw, bark, bite, lie, we, feather, seed, warm' and included the words 'near, wind, year, far, snake, short, salt, thin, heavy, worm') which is, however, not significant for our purposes. The rules of the game are formulated as follows: if two languages are genetically related, the percentage of cognates within the 35 word-list must be higher than the percentage of cognates within the 65 word-list. If we take Russian and English, for example, the percentage of cognates within the 35 word-list is 54% (19/35), while within the 65 word-list it is 25% (16/65). This is easily explainable: in cases of genetic relationships, cognates tend to cluster within the most stable part of the vocabulary, whereas in cases of fortuitous coincidence, they should be more or less evenly distributed among the whole list (and in cases of borrowings, they should rather cluster within the least stable part of the vocabulary).

This procedure was tested on a vast number of languages (Indo-European, Sino-Tibetan, Uralic, Altaic, Semitic and many others), and I must state that I do not know a *single* case where it does not work (such a case would be a widely accepted genetic relationship for which we would observe a different statistical correlation).

Now, if we take Trask's English-Basque list, we can easily see that it contains just 3 words from the 35-word-list (*odor* : *blood*, *bete* : *full*, *(j)akin* : *know*) which gives us 9%. The other 17 words give us a figure of 26% within the 65-word-list, and Yakhontov's rule is quite obviously violated. Note that this does not mean that Basque and English are unrelated (for all I know, *(j)akin* and *know* may be genuinely related on a "Proto-World" level, but we shall not be able in the nearest future to evaluate this possibility seriously); it just means that either we are dealing with loanwords (as in the case of *mendi* : *mountain*), or the absolute majority of the proposed links is due to chance.

Let us take Bengtson's Basque-Caucasian list again. Here we have, out of 19 items on the 100 word-list, 13 items belonging to the 35-word-list ('this', 'horn', 'ear', 'sun', 'thou', 'die', 'name', 'know', 'I', 'who', 'dog', 'fire', 'what') which gives us 37%, and leaves us with only 9 % matches within the 65-word-list. This certainly seems like a significant result to me. I must repeat again: I act only as an independent observer. I did not invent Yakhontov's methodology; I did not choose the words in the 35 and 65 word-lists; I did not compare Basque and North Caucasian. All the evidence is published, and absolutely everyone must arrive at the same results if he applies the same methodology. I am not saying (as Trask states) that I am *personally* impressed by these comparisons: I did not find them, Bengtson did, and I am just evaluating his evidence. Now, if this is not a "scientific way of proceeding", I sincerely do not know how to proceed.

5. My conclusions. My conclusions are generally the same. Note that my first conclusion is rather moderate: "Basque is quite possibly a member of the Sino-Caucasian family", while Trask's answer is an "unambiguous 'no'." But a serious way of dealing with it would be the following:

a) First, we formulate a hypothesis. To do that, we need the sort of evidence that Bengtson provided, the sort of criticism that Trask provided — which in fact I find very useful (in my first response, I have listed a large number of cases where I quite agree with this criticism) which allows us to narrow down the list of possible cognates and to dismiss all the cases that are effectively explained in other ways by modern Basque scholarship.

b) Second, we test the hypothesis by statistical methods. I firmly believe in the validity of this methodology — just because it works in *all* commonly accepted cases of genetic relationship. Note again that, if the results are negative, it does not immediately destroy the hypothesis: it just means that the set of cognates proposed is wrong. For all I know, a rule could operate in Pre-Proto-Basque that changed something like **bj-* into *tx-*, and Basque *txori* could be related to English *bird* (I myself, of course, think that it is absurd, but one must keep in mind that not all phonetic correspondences are necessarily obvious: they just have to be *regular*). And if the results are positive, it does not necessarily mean that we have established a genetic link — but it gives a strong support to the hypothesis and stimulates its further elaboration.

c) Third, we must find a set of phonetic correspondences that would work on the proposed set of cognates and elsewhere. These correspondences are the ultimate method of evaluating proposed cognates. Thus, if we find that the word for 'fire' is the only case of Basque *s* corresponding to PNC **c̨*, this would be a very good reason to refute this comparison; on the other hand, some quite unexpected cognates could be found. Until that is done, one actually has no right to refute any of the proposed cognates — unless, of course one finds philological flaws, wrong segmentation of morphemes and outright bad semantics in all comparisons proposed. Trask has done a very good job from the Basque side, but he failed to destroy all the comparisons. I do not regard his corrections as "minor irritants"; I would strongly advise anybody doing research on Basque's external relationships to pay attention to this criticism. But despite all efforts, the hypothesis still holds until we have a set of unrefuted cognates and no phonological means to refute them.

d) Only after establishing phonetic correspondences, shall we be able to tackle the tricky problems of Basque historical morphology outlined by Trask. I am quite sure that any problems of historical morphology can only be resolved with a good knowledge of historical phonology, and the latter is obtained only by comparing words.

I am still rather pessimistic about the future of Basque studies. Replying to this statement in my reply, Trask has obviously misread the text. By "Basque studies" I do not mean, of course, the internal Basque scholarship, which, as far as I can see, is developing quite successfully, although not devoid of blunders; I think it is quite evident from my text that I am worried about the fate of external Basque comparisons (and of long-range comparisons in general). So I shall forgive the bluntness, since I really am no specialist in Basque. I do not think, however, that specialists can dismiss other people's suggestions. As far as Basque phonology is concerned, for example, you do not have to be a specialist in Basque to see several very strange and typologically uncommon features of the proposed Proto-Basque system. Shouldn't the specialists pay attention to outsiders' views in such cases?

To quote Trask once again:

> In any discussions between specialists and comparativists, it is perfectly clear what the role of the specialists should be: to provide the most comprehensive and reliable information possible about the prehistory of their specialist languages or families.

Very well, but can all this information always be taken at face value? I have discovered several times that information provided by specialists sometimes turns out to be of very dubious quality. This, in fact, forced me to become a specialist myself in several linguistic families. But Trask seems to think that all long-rangers should just "fall joyfully on their results", without any critical evaluation. This is not the sort of cooperation we need. We shall joyfully accept your data and your corrections, but you should be prepared to modify some of your conclusions, too. For instance, there are two aspects in the Proto-Basque reconstruction that I am very skeptical of:

1) The absence of **m*. In Eurasia, at least, I do not know any languages without a labial nasal (reportedly, there are a couple of cases like that in North America, so the possibility is not entirely excluded — however, typologically rather improbable);

2) The absence of word-initial **p*, **t*, **k*. Again, I do not know of any languages with just voiced initial stops (but otherwise possessing the voice contrast).

Of course, typological arguments are not decisive in reconstruction, but I think it is widely accepted that they should be accounted for, and that a reconstructed system is more plausible if it conforms to standard typology. So shouldn't we be searching for other solutions in Basque historical phonology?

To sum up: we need cooperation, but cooperation should be mutual. Long-rangers should be prepared to accept the results of research in any particular field, and specialists should be prepared to accept reasonable propositions and corrections from the outside.

Unfortunately, among the specialists at this point, there are few who are ready for concessions, and this is the real reason for my pessimism.

MOTHER TONGUE
Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Second Response to Starostin

R. L. Trask
COGS, University of Sussex
Brighton BN1 9QH United Kingdom
larryt@cogs.susx.ac.uk

1. In his thoughtful and very interesting second response, Starostin raises several points to which I think I must respond. I shall follow his order of presentation here.

Starostin notes my pessimistic assessment of the practice of working on languages you don't know well and concludes that I am arguing that long-range comparison is impossible in principle, including the case of Nostratic. Not so.

There is more than one way of going about long-range comparison, and the Nostratic case is very different in nature from some of the other cases I have seen. Illič-Svityč worked on six families, all of them except Altaic securely established, and he made every effort to work only with the six proto-languages reconstructed by specialists in the relevant families, insofar as this information was available, as in many instances it was. I have no quarrel whatever with this procedure, and indeed I believe that it is ultimately the only valid way of establishing remote genetic links, if such links exist and are not beyond discovery. But Bengtson, and others whose work I have criticized, do not proceed in this manner at all. Instead, they simply extract items incomprehendingly from dictionaries and other secondary sources; they make no effort whatever to find out anything about the histories of those items or about the known histories or prehistories of the languages involved, and furthermore they furiously reject such information when it is offered to them. If Bengtson were trying to base his comparisons on what is known about Pre-Basque, I might have some sympathy, but he is doing no such thing, and his efforts have so far led him into just such disastrous conclusions as I have warned against.

Starostin goes on to compare long-range linguistic comparison to biological taxonomy, and asserts that my position would require the abolition of this last. Again, not so. The parallel between languages and organisms is at times illuminating, but at other times it is seriously misleading — and here we certainly have one of the misleading cases. Living creatures carry within their bodies abundant evidence about their own evolutionary history, partly in the form of highly stable structures like backbones but more particularly in the form of their genes, which change so slowly and persist for so long that biological taxonomists have a wealth of information to work with — hence the very great success of this field. But languages most definitely do *not* carry equivalent information about their own origins across countless generations. Instead, the remorseless processes of phonological change, grammatical restructuring and lexical replacement gradually but surely extinguish all evidence of the origin of every language, and hence all evidence of a possible common origin for some languages. Linguists may reasonably differ over the maximum time depth across which a common origin may be recovered, but everybody agrees that such a time depth exists (except, of course, for the Proto-Worlders, whose ideas I, like most others, dismiss as absurd beyond comment). Moreover, contact between languages provides another mechanism for obscuring ancestry — and there is no parallel to contact in the biological world. Classifying languages is *fundamentally* different from classifying organisms.

Starostin closes his first section with another plea for cooperation. Well, I think I can say I've done my best to offer my cooperation, by explaining ceaselessly what has already been established about the prehistory of Basque. Bengtson and Ruhlen have flatly refused to pay any attention to my interventions, and have preferred instead to pursue their own fantastic ideas about that prehistory. Who is guilty of a lack of cooperation?

2. Starostin concludes that no genetic link can be regarded as established until phonological correspondences have been worked out. Of course, I agree entirely, but I would add a qualification: those correspondences must involve Basque words which can reasonably be regarded as native and ancient. Be they ever so impressive, correspondences involving blatant loan words, neologisms, nursery words, arbitrarily selected chunks of longer words, grossly misanalyzed compounds and derivatives, transparently secondary regional variants, and comical semantic mismatches do not constitute evidence.

3. Starostin asserts that it is quite possible that such modern Basque words as *azeri* 'fox', *kokot* 'nape' and *xahu* 'clean' may be native and ancient. Forgive me, but to anyone who is acquainted with the history of Basque this is comparable to asserting that it is quite possible that *treasure*, *picture*, and *country* are native and ancient English words. The word *kokot* in particular looks about as much like a native Basque word as *pizza* looks like a native English word.

Starostin argues that the absence of an */m/ in Pre-Basque would not harm Bengtson's case, because Bengtson has identified a semi-regular correspondence between Caucasian */m/ and Basque */p/ or */b/. No, this is not so. As Bengtson himself has often asserted, his case crucially *requires* Pre-Basque to have had an */m/, and a number of his comparisons will fall at once if no such segment was present. If Bengtson wants to modify his case along the lines suggested by Starostin, he is free to do so, but so far he has attempted no such thing.

Now to Basque *zori* 'luck' and *(t)xori* 'bird'. Obviously, I have failed to explain myself clearly, so let me try again. The original Basque word for 'bird' was **zori.* This, in typical Basque fashion, acquired an expressive (or diminutive) variant *(t)xori.* Compare *zakur* 'dog', with its similar expressive variant *(t)xakur.* In some areas today, *zakur* is still 'dog', while *(t)xakur* means 'little dog'. Elsewhere, *(t)xakur* has become the unmarked word for 'dog', while *zakur* is now either confined to 'big dog' or lost altogether. There are dozens of such cases in Basque. In the case of the word for 'bird', the original **zori* became specialized in the transferred sense of 'omen', which has more recently developed into 'luck, fortune', and the original expressive form *(t)xori* has everywhere become the ordinary word for 'bird'. Compare Latin AVICE 'bird', which developed into Old Castilian *auçe,* at first 'bird' but then only 'luck, fortune', while the sense of 'bird' was transferred to the new lexical item *pájaro*.

I cannot see why this account should raise any eyebrows at all. In any case, it is certain that no indigenous lexical item in Basque ever contains *tx* or *x*; these segments are confined to expressive variants. Hence the comparison of modern Basque *txori* with a Caucasian word containing initial */č'/, which depends crucially upon taking Basque *tx* as original and ancient, may be dismissed out of hand. The match may be superficially beguiling, but so is the much better match between English *much* and Spanish *mucho* 'much', which are known to be unrelated. Comparing strictly modern forms is acceptable if no better information is available, but, when, as in the case of *txori*, better information is available, ignoring it is simply irresponsible.

Starostin goes on to question my assertion that Basque forms diminutives by converting other segments to *(t)x*. Perhaps, then, he would like to contemplate what happens with proper names? Spanish *Domingo* 'Dominic' forms a Basque diminutive *Txomin*; *Jose Maria* forms *Txema*; *Miguel* forms *Mitxel* (present in the surname of the great Basque linguist *Mitxelena*); *Catalina* forms *Txatalin*; *Madelena* forms *Txadalen*; *Nicolás* forms *Txekolas*; *Pedro* forms *Txeru*; *Martín* forms both *Txartin* and *Matxin*; and so on. Is that good enough? If not, how about some more ordinary words? Basque *zezen* 'bull' forms *xexen*; *zoko* 'corner' forms *(t)xoko*; *zuri* 'white' forms *(t)xuri*; *guti* 'not much' forms *gutxi*; *tiki* 'small' forms *txiki*; *tu* 'spit, saliva' forms *txu*; the Spanish loan *popa* 'poop' (of a boat) forms *txopa*; and so on, and so on. Does Starostin believe that I make these statements up out of thin air?

4. This is by far the most interesting section of Starostin's letter. Starostin adduces the work of Yakhontov, who divides the Swadesh 100-word list into two smaller lists of 35 and 65 words, and then reaches the following conclusion: if two languages are genetically related, the percentage of cognates within the 35-word list must be higher than the percentage of cognates within the 65-word list. Starostin goes on to report that this conclusion has so far proved to be exceptionless when applied to languages known to be related.

This is fascinating information, and I regret not having heard it before, or I would have mentioned it in my recent textbook of historical linguistics. It certainly looks like a substantial conclusion. But what does it have to do with Bengtson's Basque-Caucasian comparisons? Nothing. Absolutely nothing at all. Let me try to explain why.

Yakhontov's principle (and let us grant its validity) is a finding which applies to languages which have *already* been shown to be related by conventional methods, and between which secure cognates have *already* been identified by systematic correspondences. But Basque and Caucasian do *not* constitute a case of languages known to be related, and absolutely *no* cognates have been identified. Therefore, Yakhontov's principle can have nothing to say about the Basque-Caucasian case. So why on earth is Starostin adducing the principle here, where there are no cognates to talk about?

This seems a very odd thing to do, but it quickly becomes clear what Starostin is up to. He wants to invoke an entirely *different* principle, as follows: if two languages are genetically related, the percentage of *miscellaneous resemblances* within the 35-word list must be higher than the percentage of miscellaneous resemblances within the

65-word list. From this entirely new principle, he now wants to draw a *further* conclusion: if the situation just described obtains, we can count it as evidence that the languages under comparison are genetically related.

But this a shrieking *non sequitur.* Starostin's new "principles" are not at all the same thing as Yakhontov's principle, and they are supported by no evidence at all: he has merely asserted them. So why should anyone believe them?

But it's worse: Starostin's new principles are not even testable, because neither he nor anyone else has provided any definition of what is to be counted as a "resemblance". As I complained in my first response, we are left with nothing but the following: something is a resemblance if Starostin says it is. If anyone disagrees, there is no basis for rational discussion. And I can see no reason for regarding this as a "scientific way of proceeding".

Starostin goes on to point out that he did not personally find any of these comparisons; Bengtson found them, and he (Starostin) is merely evaluating them. But that makes things still worse, because Bengtson's comparisons have been carefully selected to promote the conclusion he wants to reach, and do not at all represent an unbiased assessment of the data.

Starostin finds, among Bengtson's published comparisons, 13 items from Yakhontov's 35-word list. Let's consider the first few of these.

The first is Basque *hau(r)* 'this'. But I can't find any mention at all of this word in any of Bengtson's publications, and I don't think he's ever cited it in a comparison. Of course, Chirikba has cited it, but, as I pointed out in my long MT article, Chirikba's comparison is an unholy mess: he confuses two distinct stems of different meaning, treats case-endings as parts of stems, and cites only the western form *hau* when it is clear that eastern *haur* is more conservative. If this little disaster area (wrongly attributed to Bengtson) is what Starostin is willing to count as an impressive comparison, then my previous point is vindicated at once: Starostin sees impressive resemblances where I can see only ignorance and confusion.

Next we have Basque *adar* 'horn', which is suspected of being a loan from Celtic; compare Old Irish *adarc* 'horn', of continental origin.

The third item is Basque *belarri* 'ear'. Here that lateral is crucial to the comparison, but Bengtson can't have it. He has deliberately selected just one regional variant for comparison, the one that gives him the match he needs (after he's discarded the inconvenient morph *be-*), but eastern *beharri*, *begarri* show that the lateral cannot be original, and we reconstruct **berarri,* with familiar developments of the sequence */VrVrrV/. How would Starostin respond to someone who complained that Pre-Basque **berarri* and Proto-North-Caucasian **lĕV(r)hlV* didn't look much like a match? What criterion would he appeal to defend his judgment?

With Basque *eguzki*, *eki* 'sun', we are looking at one of the most catastrophic mess-ups of all. The Basque for 'day' is *egun,* whose regular combining form is *egu-*; this combines with the noun-forming suffix *-(z)ki* to form western *eguzki,* eastern *eki,* 'sun', with regular phonology in both cases. Bengtson compares *egun* with one set of words, *eguzki* with a second (different) set of words, and *eki* with a *third* set of words! The result is a travesty, and it would be laughed away by any Vasconist at first glance — but Starostin once again pronounces himself satisfied.

I see little point in continuing with this scrutiny. Bengtson's catalogue of blunders and carefully selected regional variants has been accepted at face value by Starostin, who applies his mysterious new principle to the resulting mess, and then declares "This certainly seems like a significant result to me." To me, I'm afraid, it seems like nothing but confusion piled on confusion.

5. Starostin declares that I assert that Basque is "unambiguous[ly]" not part of the Sino-Caucasian family. Not so. It is never possible to assert that languages are not related; this is a logical impossibility. All I have ever asserted is that there is *no evidence* that Basque is related to the other languages in question — a position I am happy to reaffirm here.

I find little to object to in Starostin's summary of proper methodology, but once again I do not regard the procedure he adopts in this letter as constituting any kind of reasonable statistical procedure.

Starostin explains that, by the phrase "Basque studies", he intends, not what everybody else understands by the term, but merely the external comparison of Basque. This is a novel usage, but OK. In that case, I too share his pessimism, but probably not for the same reasons: more than a century of the most dedicated searching after a relative for Basque has failed to turn up anything of interest, and I think we are now close to the point at which we must accept that, in all likelihood, no relative for Basque will ever be discovered. There are practically no stones left unturned.

I am curious, though, to know what Starostin means by saying that we Vasconists have committed "blunders". Would he care to elaborate on this?

Starostin questions the contribution of specialists, observing that some specialists have sometimes given him poor information. Sure, no specialist is infallible, and the occasional specialist can be very pigheaded indeed. But I'm sure Starostin does not intend to imply that the conclusions of specialists are, in general, questionable or worse. Most specialists are very competent within their areas, and most of their conclusions are backed up by an intimate knowledge of the evidence which is not available to outsiders. Starostin's position here is reasonable, but it contrasts sharply with the views of some other long-rangers, who will listen to nothing said by specialists unless it suits their own narrow purposes, and who will unblinkingly wave away any amount of counterevidence provided by specialists if it gets in the way of their precious comparisons.

Starostin goes on to query two aspects of the standard reconstruction of Pre-Basque phonology, both of which trouble him.

First, he dislikes the absence of */m/, and he asserts that languages without /m/ are very rare and perhaps unknown in Eurasia. Well, such languages are admittedly not thick on the ground, but they're not exactly rare, either. I come from western New York State, where the indigenous language is the Iroquoian language Seneca. I only have to drive down the road a bit to hear Seneca spoken — and Seneca has no /m/, perhaps uniquely among Iroquoian languages. Among other languages lacking /m/ are Wapishana (Arawakan), Barasano (Tucanoan), Siriono (Tupi-Guaraní), Apinaye (Ge), Mura (Paezan), Dakota (Siouan), Puget Sound (Almosan, I think), Quileute (Chimakuan), Mixtec (Oto-Manguean), Navaho (Athabaskan), Rotokas (Papuan), Hakka (Burmic), Kpelle (Mande), Lusootsheed (Salish), and possibly Cherokee (Iroquoian). This list is by no means exhaustive; it's merely representative. One or two of these languages present difficulties of analysis, notably Rotokas, in which [b] and [m] are in complementary distribution, and we have to decide what the phoneme is. The UPSID database takes the phoneme as /b/, and classifies Rotokas as having no /m/. The Papuan specialist William Foley opts for /m/, but he does so explicitly on typological grounds — he doesn't like having a language without /m/ — and so he is begging the question.

Languages without /m/ are far from rare, but, in any case, typological considerations cannot be allowed to override hard data. We cannot find any evidence for an */m/ in Pre-Basque, and we are not going to posit a gratuitous */m/ just to soothe fevered typological brows or just to make dubious comparisons look better.

Second, Starostin complains that, in a language with a voicing contrast in plosives, it is unheard-of for only the voiced ones to be permitted in word-initial position. Maybe so, but this is not relevant. One of the central points of Michelena's reconstruction of Pre-Basque is that the language *had no voicing contrasts*. We use the symbols */p t k b d g/ merely for orthographical convenience, and to remind ourselves of the modern reflexes of these six segments, but we emphatically do not interpret these symbols as representing phonetic [p t k b d g]. The contrast between */p t k/ on the one hand and */b d g/ on the other was entirely confined to word-medial position and mostly to intervocalic position; elsewhere it was neutralized — and the only plosives permitted in initial position were */b g/ and, in rare circumstances, */d/, the segments which have developed into the modern voiced plosives. Michelena characterizes the two sets of plosives as "fortis" and "lenis"; I have suggested that the fortis ones might have been no more than geminates. Either way, Starostin's point is of no relevance to Pre-Basque, because Pre-Basque didn't have any voicing contrasts.

Starostin asks: "shouldn't we be searching for other solutions in Basque historical phonology?" I reply: why? Our standard reconstruction accounts magnificently for the facts of Basque, which is all that we can reasonably expect of it. It accounts for the forms of native words, including the attested dialect variants; it accounts for the forms of loan words; it accounts for the forms of compounds and derivatives; it largely accounts for the peculiar distribution of the aspiration. What is left to be explained? Very little that I can see, apart from the word-accent, which is still a headache. We can see no problems with our reconstruction, and no other reconstruction has ever been proposed, nor is it easy to see how any other reconstruction could account equally well for the facts. Of course, it's clear that our reconstruction isn't very convenient for the eager long-rangers who want to drag Basque, kicking and screaming, into some grand scheme of their own devising, but so what? Why should we worry about this? It's not our problem if Basque doesn't fit very well into these schemes, and we won't lose any sleep over it. And we are definitely not going to throw away our magnificently successful reconstruction just so that optimistic longrangers can have a field day comparing anything they like.

Anybody who doesn't like our reconstruction is free to propose a different one, but this hasn't happened yet, and, if such a proposal isn't every bit as detailed and explicit as Michelena's, and if it doesn't account at least equally well for all the facts of Basque, we're not going to pay much attention to it. Why should we?

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Response of Larry Trask to Merritt Ruhlen

The electronic wizards having failed, Hal Fleming has kindly sent me a hard copy of *Mother Tongue* 26, which contains your brief response to my summing up in the big issue of *Mother Tongue*. I thought I would reply in turn with a few comments. (I've already replied to Bengtson, who sent me a preprint.)

In fact, it is very difficult to see any common ground for discussion at all. You clearly believe to be true a number of things which I, like historical linguists generally, believe to be false, and these differences, concerning such things as what can be counted as evidence, are so fundamental as perhaps to remove any possibility of discourse. Nevertheless, I'll try.

In your response, you enumerate four points, but you also make at least one further point which is not enumerated. I'll try to address these in turn.

(1) You again complain that I omit the Sino-Tibetan and Na-Dene data from my paper, and insist that this omission is destructive of my case. I can only reiterate what I said before. I am not trying to destroy the entire Dene-Caucasian edifice; I am merely objecting to the manner in which the Basque data have been manipulated and misused in trying to construct a case that Basque should be added to the Dene-Caucasian construct, and I am denying as a consequence that any substantial case can be made for relating Basque to *any* other languages at all on the basis of such evidence. Citing the additional data would have added perhaps twenty pages to an already unreasonably long paper, and would have contributed nothing useful to the discussion. If I can show that a Basque word cited was not in the language 2000 years ago, or that it was present but did not at the time have the modern form adduced in support of comparisons, or that it is a nursery word or imitative word which should not be invoked in comparisons as a matter of principle, or that the Basque word has been wrongly analyzed, or that the analysis offered is so capricious, *ad hoc,* and implausible as to render *any* comparison devoid of value, then that is the end of the matter, and adding any number of further languages to the comparison is a worthless waste of time.

As for your repeated brandishing of Basque *gose* 'hungry' and *gorotz* 'dung', see below.

(2) You assert that, if adding more languages to a comparison produces additional instances of resemblances to Basque words, this must be evidence for something. But this is simply not true. Adding more languages to *any* comparison, however ill-founded, will *always* produce further striking resemblances, and this result means precisely nothing, except that the laws of probability are not taking the day off.

In particular, you again point to Basque *odol* 'blood' and Proto-Athabaskan **dehl* (= **deɬ*) 'blood', and declare that this resemblance must mean something. But it doesn't. Any two arbitrary languages are going to exhibit some striking resemblances in form and meaning, just by sheer chance, and this one is no more impressive than my mischievous list of comparisons between Hawaiian and ancient Greek. It's scarcely even worthy of comment, especially since no such word is found, apparently, in any of the other languages you want to assign to Dene-Caucasian. Except, of course, for a few words that look vaguely similar but don't mean 'blood' — am I supposed to take these seriously?

You try the same thing with Basque *gose* 'hungry'. OK; Proto-Caucasian has something similar, *if* we accept the reconstruction on offer — and the Caucasianists of my acquaintance definitely do not accept that a North Caucasian genetic unity has been demonstrated. But let's say it's good. Now, why should this resemblance be any more interesting than Hawaiian and ancient Greek *meli* 'honey'? Well, you tell me, because something similar is also found in Hruso and in Galice. Well, tremendous. There are some 300 Sino-Tibetan languages, and exactly *one* of them, you tell me, has a similar-looking word. Wow. And there are thirty-odd Na-Dene languages, and exactly one of these, too, has something similar. Double wow.

Just as I've been insisting all along, every time you throw another bunch of languages on the pile, you find some more hits. So, for example, you toss on 300 Sino-Tibetan languages, and you get 299 misses and one hit for Basque *gose.* But you attach zero weight to those 299 misses and claim that the one hit is deeply significant. This means that you are claiming that 299 languages have lost the Proto-Sino-Tibetan word for 'hungry', and that Hruso alone has retained it. And why should any sane person believe that?

Before you can compare a Hruso word with a Basque word, you must make a reasonable case that the Hruso word is an archaism retained from Proto-Sino-Tibetan, something which you not only have not attempted but apparently do not even see the need for. How you can continue to deny this elementary point of logic simply mystifies me. You might as well try to make a case that the same people built ancient Rome and ancient Kyoto on the basis of one building in Kyoto which looks a bit like the Temple of Minerva but is not known to have existed before the 20th century.

You go on to ask "why all of these 'accidental' resemblances constantly fall among the **same set of families,** rather than a different set of families". The answer is simple: they don't.

No one who has ever searched for resemblances between Basque and any other language(s) at all has failed to find some, and the length of each such list represents nothing more than the amount of time and effort devoted to compiling it (with a modest qualification for the effect of phonological systems: languages which have similar phonological systems are more likely to exhibit chance resemblances than those that don't). I myself have had no trouble in compiling impressive-looking lists of "cognates" linking Basque to English and to Hungarian. But I was just playing. What about the serious efforts?

Some people have compared Basque with Afro-Asiatic, using the same methodology you use. This project was begun by Schuchardt, but the main man here is Mukarovsky. Mukarovsky spent years searching for Basque-Afro-Asiatic matches, and of course he found them in profusion. In fact, he concentrated almost entirely on Basque and Berber, with only occasional glances at other branches of Afro-Asiatic, and Berber is only one language. But, even for Berber alone, he found a long list of hits — I don't know how long, because I don't have copies of all his papers, but he found lots of matches for lexical items, grammatical words, and bound morphemes, and he published a list of phonological correspondences supposedly linking the two languages. His single most striking match is this one: in both Basque and Berber, the sex of a singular male addressee is marked with a verbal suffix *-k*, while a female addressee is marked with *-m* in Berber but *-n* in Basque (all nasals are neutralized as /n/ in word-final position in Basque). Pretty good, eh? But Afro-Asiatic is not supposed to belong to Dene-Caucasian, so *how do you explain Mukarovsky's results*?

For me, and for most of us, the answer is simple: just like you and Bengtson, Mukarovsky is doing nothing but collecting lists of miscellaneous chance resemblances of the sort that can *always* be found. But you don't buy that, so what's your explanation?

Bear in mind that, if you and your friends had found that *-k* and *-n/-m* sex marking in Burushaski or Yeniseian, you'd be brandishing it triumphantly as evidence that that blinkered reactionary Trask was blind to the obvious truth. Wouldn't you? So why don't you likewise accept Mukarovsky's case that Basque is related to Berber, and hence (in your scheme of things) a member of Nostratic, rather than of Dene-Caucasian? I await your answer with interest.

Or take Kartvelian. Trombetti, Dumézil, Bouda, and Lafon (at least) proposed lists of correspondences between Basque and Kartvelian; the resulting list is lengthy and includes pronouns, grammatical words, and bound morphemes, and Lafon even published a list of phonological correspondences. But Kartvelian isn't supposed to be Dene-Caucasian either, so, once again, how do you explain these results? And why don't you take them as further confirmation that Basque belongs to Nostratic, or Eurasiatic, or whatever?

According to your lights, the case for relating Basque to Afro-Asiatic and to Kartvelian, and hence for assigning it to Nostratic, is a perfectly sound one. But, insofar as anything is inconsistent with anything in your scheme of things, this result is inconsistent with your declared view that Basque must belong to Dene-Caucasian. So, on what basis do you buy the Basque-Dene-Caucasian case and reject the Basque-Nostratic case? Why should I believe you and not Mukarovsky, Lafon, and company? Why should I believe any of you? Can you give me a good reason why I should prefer the Dene-Caucasian case to the Nostratic case? For your case, you must realize, this is an absolutely crucial point, but one to which you have devoted no attention at all. It's not up to me to come up with a Basque-Austric straw man to set beside your Dene-Caucasian case; a competing case is already in place, so what are you doing about it?

You and Bengtson have spent years comparing Basque with North Caucasian, with Burushaski, with Sumerian, with Yeniseian, with Sino-Tibetan, and with Na-Dene, and in every case you find lots of matches — though, of course, most of these matches are only with this family or that, and many of them are only with this *language* or that. Others have spent years comparing Basque with Afro-Asiatic, or with Kartvelian, and they too find lots of matches. To the best of my knowledge, *nobody* has ever spent years comparing Basque with any other particular languages or families and failed to find lots of matches. I submit, therefore, that the following proposition is patently true: if you spend years comparing Basque with *anything,* using the methodology that you and Bengtson espouse, you will *always* find lots of matches. Nobody has failed yet, you know.

The only sensible conclusion to be drawn from these observations is that your methodology is indefensible and that your results are devoid of value.

Apart from one or two obvious cranks, nobody I know of has ever tried seriously to relate Basque to Indo-European, and I'll bet you and Bengtson haven't, either. But why not? It would work, you know. An exasperated Michelena, in reviewing the work of an earlier long-ranger, once spent five minutes rattling off an impressive list of Basque-Indo-European matches. And I can see no reason to doubt that somebody who spent five years on such a project could demonstrate that Basque is just as closely related to Indo-European as it is to Caucasian or Burushaski. And that would be deeply embarrassing for you, wouldn't it?

While I'm talking about methodology, have you seen Lyle Campbell's recent brilliant demonstration that Old Japanese is an Amerind language? If not, I'll be happy to send you a copy. Not only does Campbell demonstrate that Old Japanese fits magnificently into the Amerind etymologies proposed by Greenberg, he shows that it fits into *more* Amerind etymologies than does any single North or South American language. By your criteria, then, Japanese belongs to Amerind, and that's the end of it. Are you happy?

(3) You complain that I dismiss as accidents the 68 comparisons that I have been unable to remove from the Basque side alone. So I do, and with very good reason, as I have just explained in the preceding section. But there are further reasons for dismissing them. For one thing, these remaining comparisons do not involve all the languages under discussion, or even most of them: they are merely comparisons with this language or that, or occasionally with a couple of others. For another, the specialists in Burushaski and Sumerian have already made it clear that the comparisons involving those languages suffer from the same fatal errors as the Basque data, and hence that a further unspecified, but certainly large, number of additional comparisons must be discarded as unsustainable. No specialist in Sino-Tibetan or Na-Dene has been heard from, but I wonder what such a specialist might make of the comparisons involving those families, especially when you pick a word from one single language and happily project it back 15,000 years into the past (or whatever your time depth is).

You describe my position as "a bizarre twist on normal comparative linguistics". It is nothing of the sort. Where on earth do you get the idea that a common origin is always preferred to accidental resemblances, by default, in evaluating miscellaneous similarities between languages? This is blatantly not true. Historical linguists who have written on methodology have constantly drawn attention to the problem of chance resemblances and the need to eliminate them by the use of rigorous criteria; see for example what Anttila and Hock have to say about the matter in their textbooks. If you adopt the opposite position, you may get lucky once in a while, but mostly you will simply doom yourself to announcing an endless series of "genetic links" which are totally spurious — an outcome which I think you and your colleagues have demonstrated more than abundantly.

I have not seen the Sarich[1] article you refer to, though I have both the article and the book containing it on order; I may have some further comments after I've seen the piece. But I note that Sarich is an evolutionary biochemist with no experience of doing historical linguistics, so I'm curious to know just what insights he's offering us into our trade. I certainly wouldn't be comfortable offering evolutionary biochemists advice about *their* business, even though I have a substantial background in chemistry.

(4) I am not responsible for Goddard and Campbell's statements, and I can't comment, except to say that I am pleased to be mentioned in the same breath as two such distinguished scholars. As for "Austro-Basque", sure, I'm confident I could make a good case. But I'm not mad enough to devote five years of my life to such a pointless exercise.

It is not up to me to "prove" that your Basque-Dene-Caucasian hypothesis is wrong, and proving a negative is logically impossible anyway. It's up to you to support your hypothesis with a case that is strong enough to be taken seriously and to stand up to scrutiny, and so far you haven't done that, nor do I think it likely that you can. Anyway, the Basque-Afro-Asiatic and Basque-Kartvelian people mentioned above have already done the job for me: they have proved that, with the use of mass comparison, Basque can be shown to be related to any damn thing at all. Until you have a good explanation for these other results, you are in no position to jeer at me because I can't be bothered to add one more hopeful comparison to the inventory.

---

[1] Sarich, Vincent M. 1994. "Occam's Razor and Historical Linguistics", in: *In Honor of William S.-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. By M. Y. Chen and O. J. L. Tzeng, pp. 409-430. Taipei: Pyramid Press. [Editor]

(5) Finally, you have once again trotted out Basque *gorotz* 'dung', this time comparing me in the process to O. J. Simpson's lawyers, and you twice accuse me of "manufacturing evidence". Really, Dr. Ruhlen. First of all, my remark about Tibetan and Apache is rather obviously a piece of rhetorical speaking. The only point of substance is whether *gorotz* is a native Basque word or not. The consensus of specialists at present is that it probably is not, and this conclusion is reinforced by the account provided by Corominas (who has no stake in the matter, and has taken no position on whether Basque is or is not related to anything else).

We think *gorotz* is a loan word. Let's briefly see why.

Let's divide Basque words into two groups: those which are shared with neighboring languages, and those which are not. Of the ones which are not shared, a number must be regarded as late formations and removed from consideration. Of those still remaining, the great majority are accepted by all specialists as native Basque words. Of these native words, the majority begin with a vowel, but quite a few begin with plosives. With just a tiny handful of exceptions, most of which have good explanations, the only initial plosives found are /b/ and /g/. We know why this is: /b/ and /g/ were the only permitted word-initial plosives in the Pre-Basque of 2000 years ago. There are quite a few of these; here's a list off the top of my head:

biga 'two', *gu* 'we', *buru* 'head', *gau* 'night', *belarri* 'ear', *gatz* 'salt', *bizar* 'beard', *gero* 'later, after', *bihotz* 'heart', *gaitz* 'difficult', *behaztun* 'bile', *gerri* 'waist', *bizkar* 'back', *gorri* 'red', *berdin* 'same', *galde* 'question', *bero* 'hot', *gazte* 'young', *bigun* 'soft', *giro* 'atmosphere', *bide* 'road', *gorroto* 'hatred', *begi* 'eye', *gurin* 'animal fat, butter', *bazter* 'edge, corner', *gain* 'top', *barru* 'interior', *gai* 'material', *bortz* 'five', *goi* 'height', *bat* 'one', *garai* 'high', *barda* 'last night', *giltza* 'key', *bare* 'slug; spleen', *gibel* 'liver', *beltz* 'black', *giltzurrin* 'kidney', *ben* 'serious', *gorosti* 'holly', *berri* 'new', *goroldio* 'moss', *bihi* 'grain', *gose* 'hungry', *buztan* 'tail', *gari* 'wheat', *burdina* 'iron', *guzti* 'all', *behi* 'cow', *guti* 'not much', *baso* 'woods', *gogo* 'soul, mind', *bizi* 'alive', *gogor* 'hard', *bihar* 'tomorrow', *gehi-* 'amount', *behar* 'necessity', *gurdi* 'cart', *barre* 'smile', *gar* 'flame', *bazkari* 'lunch', *garbi* 'clean', *behor* 'mare', *gizen* 'fat', *bular* 'breast, chest', *gizon* 'man', *bertze* 'other', *gudu* 'combat', *berun* 'lead' (metal), *gorro* 'phlegm', *bera* 'soft', *guren* 'edge', *beti* 'always', *gor* 'deaf', *belar* 'grass', *guri* 'fresh, soft, tender', *borda* 'hut, shed', *gozo* 'sweet, tasty', *bider* 'occasion', *garagar* 'barley', *birika* 'lung', *garri* 'affliction, discomfort', *bele* 'crow, raven', *gordin* 'raw, crude', *beso* 'arm', *gaztai* 'cheese', *buztarri* 'yoke', *garo* 'fern', *bederatzi* 'nine', *goiz* 'morning', *bete* 'full'.

This list is far from exhaustive, but it'll do. Now, why have I bothered to list all these? Because these native words have a very interesting property: they do **not** show regional variants with initial voiceless plosives. *Nowhere* do we find **peltz* for 'black', or **kose* for 'hungry', and so on for all the others. There is just one marginal exception: the word *gar* 'flame' is attested as *kar* in one corner of the French Basque Country. But that's it: the whole rest of the country knows only *gar*, and not one of the other words is even marginally attested with a voiceless plosive. This, of course, is exactly what we expect, given the established prehistory of Basque phonology.

But now let's look at the other list, the list of words shared with neighboring languages. This list contains some words attested only with initial voiced plosives, such as *diru* 'money', borrowed from Latin DENARIU, and some attested only with initial voiceless plosives, such as *kaiku* 'wooden bowl', from Latin CAUCU. Crucially, however, it contains a large number of items which are attested with *both* voiced and voiceless plosives in large areas of the country. Here are just a few:

biko ~ piko 'fig', *bake ~ pake* 'peace', *bago ~ pago* 'beech', *biper ~ piper* 'pepper', *boltsa ~ poltsa* 'handbag, purse', *berna ~ perna* 'leg', *berma ~ perma* 'loss', *berde ~ perde* 'green', *bekatu ~ pekatu* 'sin'.

dekuma ~ tekuma 'tithe', *dipula ~ tipula* 'onion', *dolare ~ tolare* 'wine press', *dorre ~ torre* 'tower', *dorpe ~ torpe* 'heavy, sluggish'.

gatea ~ katea 'chain', *gatu ~ katu* 'cat', *gapelu ~ kapelu* 'cap', *gereta ~ kereta* 'rustic gate', *golko ~ kolko* 'bay, gulf', *goro ~ koro* 'vault, dome', *gorta ~ korta* 'courtyard, town common', *gutizia ~ kutizia* 'desire', *gutun ~ kutun* 'letter', *gako ~ kako* 'hook', *grina ~ krina* 'passion', *ganbera ~ kanbera* 'room, chamber', *gapirio ~ kapirio* 'stripe'.

Every one of these words is shared with neighboring languages, and every one is universally accepted as a loan word from one of those neighboring languages; if you know a bit of Latin and Romance, the sources of most should be obvious. Now, why the fluctuation in the voicing of the initial consonant? Well, with just one exception (the word for 'tithe'), the word in the original language had an initial voiceless plosive. But Pre-Basque didn't permit initial voiceless plosives, and so the words were generally borrowed with initial voiced plosives. At some stage, however, these words acquired initial voiceless plosives as permitted variants. It's not clear why this happened — maybe the words were later reborrowed, or re-formed under the continuing influence of neighboring Romance languages which retained the voiceless initial, or maybe there were just different strategies for dealing with such loans, depending on the time and the place. But this doesn't matter. The key point is this: fluctuation in the voicing of initial plosives is highly characteristic of loan words, but unknown in native words.

And now you'll see the point: the word *gorotz* 'dung' has an equally frequent variant *korotz*, and both forms are found throughout the country — that is, neither form is only marginally attested, or confined to a small area. Indeed, the two forms seem to exist side by side in every corner of the country.

So, what conclusion can we draw? It's obvious: *gorotz* ~ *korotz* must be a loan word, even if it's a loan word whose source is not entirely certain. It must be a loan word, because it *behaves* like a loan word, and it can't be a native word, because it doesn't behave like a native word. Therefore, it must be removed from consideration in any comparison of Basque with other languages, and hence any number of impressive-looking Dene-Caucasian comparanda involving this word are worthless as evidence for a genetic link between Basque and Dene-Caucasian.

This is the sort of thing that we blinkered reactionaries count as evidence. Now, if you still want to stick to your guns, you are forced to conclude that this one Basque word, so crucial to your case, is absolutely unique among native Basque words in behaving like a loan word, instead of like a native word. If you still want to do this, I refer you to that very Occam's razor cited in the title of the Sarich article that you so obviously approve of: if it looks like a duck, and walks like a duck, and quacks like a duck, then it's a duck, and any argument that we are looking at the only chicken in the world that behaves like a duck must be dismissed.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Letter from Merritt Ruhlen to Larry Trask

Thanks for your letter of July 17. I'll try to respond to the questions you raise.

Let me begin by saying that I think you — and historical linguists in general — simply misunderstand the nature of the comparative method. The comparative method consists of two largely independent steps. The first step is taxonomy, classification, or multilateral comparison, whichever term you prefer. By classifying languages on the basis of shared lexical and grammatical similarities, one identifies language families — at all levels. This is how Indo-European was discovered, and it is in fact how all families are discovered. There really is no other way. The second step in the comparative method is what has traditionally been called Historical Linguistics, and this stage involves the reconstruction of the proto-language, discovery of sound laws, determination of a homeland, and all the other problems raised by any linguistic family. The oft-cited claim by you and other historical linguists that it is reconstruction and sound laws that "prove" a language family is not only absurd, it is also strictly a twentieth century myth, which is probably why Hal Fleming refers to it affectionately as "Indo-Baloney". One finds no mention of this myth among the nineteenth century pioneers of Indo-European. The difference between these two steps in the comparative method is so obvious it is hard to see how they have become so confused in the minds of contemporary historical linguists. One need only recognize that Joseph Greenberg, in the past half century, has essentially worked out the genetic relations of African, New Guinean, and American languages by means of multilateral comparison, while critics of this method have made essentially no contribution to taxonomy throughout the entire twentieth century. Can you think of one contribution the Indo-Europeanists — or the Americanists — have made to linguistic taxonomy this century? Neither can I. And the reason is very simple. They don't do taxonomy and therefore they will never find any new genetic connections. Waiting for Goddard, Campbell, or Watkins to find a new linguistic relationship is like waiting for Godot.

In light of the above, it should be clear that one does not "add" Basque to a family, but rather one classifies all relevant languages and sees where Basque fits in. This is exactly what Bengtson has done, and he has presented a very strong case — following precursors of course — that Basque is closest to the other members of Dene-Caucasian. Of course, a number of the pioneering etymologies are not 100% correct, and you have provided a service in weeding out certain fallacious analyses. But what is left untouched by your critique is more than ample to demonstrate that Basque is a member of Dene-Caucasian.

Since you seem to think that one can find anything anywhere, and that you could provide just as good evidence for Austro-Basque as Bengtson has for Dene-Caucasian, it is difficult to see where language families come from. If everything exists everywhere, how could one ever detect language families? I know that yours is a widespread belief among specialists who spend their whole life studying one language, but for those of us who spend our lives studying many languages, we know that it is patently false. Wanting to find some specific form in a language, or language family, is no guarantee of finding it. Consider the N/M Amerind pronoun pattern that Campbell and Goddard claimed was found in the Old World just as much as in the New World, or the M/T Eurasiatic pronoun pattern that Campbell and Goddard claimed was just as prevalent in the Americas as in northern Eurasia. I believe I presented definitive evidence in *Mother Tongue* that M/T is virtually non-existent in the Americas, while N/M is all over the place in both North and South America. That Goddard and Campbell could produce no evidence to back up their claims is eloquent testimony to the fact that not everything exists everywhere. So your similar claims, without a shred of evidence, cannot really be taken any more seriously.

I'm afraid I can't accept your plea that you didn't have room for Sino-Tibetan or Na-Dene. The title of your paper says you are considering all of Dene-Caucasian, not just part of it, and it is moreover methodologically unsound to compare any language with *part* of a family. And consideration of all of Dene-Caucasian does provide additional evidence for the family as a whole, as Na-Dene **deɬ* 'blood' and **gase* 'hungry' attest. And what about Basque *agor* 'dry,' which has cognate forms in every branch of Dene-Caucasian? I know, just another accident.

You claim that as one adds more and more languages to the comparison, one will inevitably find more and more accidental resemblances. Of course. But what one won't find — if the resemblances are truly accidental — is that these accidental resemblances consistently fall within a well defined set of languages. If they do, and clearly they do for Dene-Caucasian, then they cannot be accidental. Accidents do not pattern themselves regularly. This is obvious to those of us who work with many languages; it's really only narrow specialists who harbor the illusion that accidental resemblances occur all over the place. They really don't. Accidents, by their very nature, are few and far between. If **deɬ* 'blood' and **gase* 'hungry' were the only two similarities between Basque and Na-Dene, one might well consider them accidental. But they are just two of many such items that just happen to characterize Dene-Caucasian, and therefore there is very little chance that such resemblances are accidental.

By the way, **gase* 'hungry' is not found in just one Na-Dene language. It characterizes the Pacific Coast subgroup of this family, and this subgroup is in all likelihood primary subgroup of Athabaskan opposed to all the rest (note in the Na-Dene paper I sent you that several items are restricted to Pacific Coast Athabaskan). With regard to the one Sino-Tibetan language cited, Hruso, you complain that one language doesn't mean anything. Unfortunately we don't have a complete Sino-Tibetan comparative dictionary[1], so whether this particular root is found in just this one language or many others is not really known. I suspect that this root is probably found elsewhere in Sino-Tibetan (after all, since you can find anything anywhere, you would have no problem finding this root throughout Sino-Tibetan!). But we will have to await a comprehensive Sino-Tibetan dictionary — or someone will have to plow through many dictionaries of Sino-Tibetan languages — to find out. By citing just one language, Bengtson calls attention to the fact that this root may exist as well in Sino-Tibetan, and other scholars may cite additional languages if they are aware of them. Very little of Bengtson's work involves a single language from some family, as you know well, so citing this one Hruso example is kind of a red herring. Dene-Caucasian is not built on one word in one language.

You seem to put great store in the fact that "no one who has ever searched for resemblances between Basque and any other language(s) at all has failed to find some, and the length of each such list represents nothing more than the amount of time and effort devoted to compiling it." Since Mukarovsky has compared Basque with Afro-Asiatic, this disproves Dene-Caucasian? Gimme a break. I don't have Mukarovsky's articles right now, but I have seen them. First of all, as you point out, his comparisons are really more with Berber than Afro-Asiatic, and the comparison of a language with part of a family is methodologically unsound, as I pointed out above. If Basque is really related to Afro-Asiatic, it should show similarities not just with Berber but with the other branches of Afro-Asiatic as well. As I recall, there is very little of that in Mukarovsky's articles. Furthermore, I believe that the semantics and phonetics of his comparisons are not in the same ballpark as the comparisons of Bengtson and the other Dene-Caucasianists. In a word, Mukarovsky is simply wrong, and Bengtson et al. are right. One man's bad work does not undermine the good work of another.

Japanese is very similar to Basque in its taxonomic history, being considered a language isolate with no known relatives. Japanese too has been compared with everything under the sun and extensive hypotheses relating it to Tamil, Austronesian, Sumerian, Altaic, etc. have been presented. But, as in the case of Basque, the nearest linguistic relatives of Japanese are no mystery at all. Roy Miller, Joe Greenberg, Sergei Starostin, and the Nostraticists in general, all agree that it is the Altaic connection that is correct, and the evidence they have published with regard to this connection is abundant. What about the other proposals? They are for the most part absurd, like Mukarovsky's Basque proposal. Benedict's recent proposal connecting Japanese most closely with Austronesian is ridiculous and has been demolished by Miller. With regard to the Tamil proposal, it is difficult to imagine why one would compare Japanese with one Dravidian language and not the whole family. And as for Japanese-Sumerian, the less said the better. Note also that the truly absurd proposals are invariably the work of *one* scholar, usually a specialist in one of the two languages being compared, who has chosen the other language for comparison on an ad hoc basis. While there are many notable scholars who agree that Japanese is closest to Altaic, all the other proposals have a single exponent. The same is true of Dene-Caucasian. Many people support this particular theory; only Mukarovsky supports the Berber-Basque connection so far as I know. Other recent examples of such arbitrary binary comparisons that reach absurd conclusions with absurd evidence are Colarusso's attempt to connect Indo-European and North Caucasian, and Pulleybank's attempt to connect Indo-European and Sino-Tibetan. Do these proposals undermine the solid work of the Nostraticists and Greenberg that demonstrate conclusively that Indo-European's closest relatives are spread across northern Eurasia, extending into the Americas with Eskimo-Aleut? Not in the least. I haven't seen Lyle's paper showing that Old Japanese, and presumably modem Japanese as well, is an Amerind language. I would suggest you take a look at some of Roy Miller's work, either his

---

[1] I have been informed that an etymological dictionary of Sino-Tibetan by Ilia Peiros and Sergei A. Starostin has been completed and distribution was begun at the Sino-Tibetan conference in Leiden, October 1996 — it is entitled *A Comparative Vocabulary of Five Sino-Tibetan Languages* and is published by the Department of Linguistics, The University of Melbourne (1996). [Editor]

1972 book, *Japanese and the Other Altaic Languages,* or his recent article in the Rice Conference volume, *Sprung From Some Common Source,* edited by Lamb and Mitchell, Stanford University Press, 1991.

So your claim that Mukarovsky has already produced a case for Basque origins that is as well supported as that of Dene-Caucasian is not true in my opinion. Nor will you ever — or could you ever — produce an equally compelling case for Austro-Basque, connecting Basque with Khoisan, Australian, Gilyak, Algonquian, and Quechuan. The evidence is just not there. We know. We've looked.

One point of correction: You say that Bengtson and I have spent years comparing Basque with North Caucasian. Bengtson probably has, but I have concerned myself very little with any of Dene-Caucasian up to the paper I'm working on now. So all of the credit for Dene-Caucasian — or blame in your eyes — belongs to Bengtson, Blažek, Starostin, Nikolaev, and Chirikba. I am a supporter of the Dene-Caucasian hypothesis, but I am not one of its architects.

With regard to your statement that "the only sensible conclusion to be drawn from these observations is that your methodology is indefensible and that your results are devoid of value," I would suggest talking to an Africanist, for whom Greenberg's classification is the basis of all African historical linguistics. The same is true of Amerind, though the specialists haven't realized this yet. And the same is true for Dene-Caucasian, though specialists haven't realized this either — for the same reasons.

About 20 years ago — long before I had any interest in or knowledge about genetic classification — Greenberg remarked to me one day that "if you want to know where Basque fits into the world's languages, the last person to ask is a Basque specialist. He will tell you (with complete candor) that he has never seen any language like Basque. But he has also never looked at any other language but Basque. If you want to know where Basque fits into the world's languages, you should ask someone who knows *nothing* about Basque, but everything about the world's other languages." The recent debate in *Mother Tongue* seems to have shown Greenberg right once again.

With regard to your claim that common origin is not the default explanation for a set of similarities among a certain group of languages and that accidental similarities are just as likely, I'll let you read Sarich's paper[2]. You seem skeptical that you could learn anything about historical linguistics from an evolutionary biologist. Let's find out.

Let us turn now to Basque *gorotz* 'dung,' which you spend so much time on in trying to do away with Bengtson's etymology connecting this form with Proto-Caucasian **k'urč'V* 'dung' and Burushaski *γurAš* 'dung.' We have here two etymological explanations for the Basque word: (1) borrowing and (2) common origin. You seem to have changed your borrowing explanation from your article, in which you claim the word is a borrowing of Latin *crocea* 'saffron-colored,' to your letter, in which you now see this root as a borrowing from some unknown language because the initial consonant shows alternation of voicing, a trait of loan words. Bengtson's explanation of common origin from Macro-Caucasian is completely straightforward, and the fact that the semantics are precise and the phonetics agree in all three consonants is truly impressive. Sure accidents occur, but take my word that accidents involving three consonants are rare indeed. Can you think of any? It seems to me that in weighing these two explanations, Bengtson's is overwhelmingly more likely than yours. In fact, your explanation is not really an explanation at all since it involves a mysterious unknown language that loans Basque *gorotz* and then disappears without a trace.

Let me also say that I cannot accept your claim that "my remark about Tibetan and Apache is rather obviously a piece of rhetorical speaking." What you said was: "A fine example of this is Basque *gorotz* 'dung,' which is explicitly singled out by Ruhlen as a case in which the data from Sino-Tibetan and Na-Dene would prop up a comparison." No one could read this without assuming that I had supplied putative cognates from Sino-Tibetan and Na-Dene. There is nothing rhetorical about it. It is a willful deception on your part. You could have chosen an example where I *did* provide evidence from Na-Dene, such as 'blood', 'hunger', or 'dry'. But of course this wouldn't have fit your agenda, would it?

I have also sent you under separate cover Sarich's paper from the Wang volume, my paper from the Wang volume, and a paper I'm currently working on which attempts to show that within Dene-Caucasian, Yeniseian and Na-Dene share a special relationship, i.e., they form a valid subgroup within the larger family. I also enclose a recent paper of mine from a conference in Tokyo in 1992.

---

[2] See note to Trask's letter. [Editor]

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

Comments on Sergei Nikolaev's Letter To Harold C. Fleming (2 September 1995) [*Mother Tongue* (Newsletter) 26:19-22]

John D. Bengtson
Minneapolis, MN

First I should say that I think Sergei Nikolaev is one of the very best of the pioneer long-rangers. Now, a decade after his discovery (or re-discovery) of the Dene-Caucasic hypothesis (Caucasic + Na-Dene), the idea almost seems commonplace, at least within the long-ranger community. Just think — to envision a connection between these two families, separated by vast distances and racial and cultural barriers, is hardly an obvious choice. But Sergei Nikolaev, with his paper "Sino-Caucasian Languages in America" (1991) did it. He established the genetic connection between Caucasic and Na-Dene, by means of basic lexical comparisons, and further supported that claim by showing that there are systematic correspondences between the phonological structures of both families.

In retrospect, we can imagine any number of other linguists *could have* done the same — could have — but they *did not*. Others *might* have revived the comparison of Basque and Caucasic, but only Vjačeslav Chirikba did. Others who were more learned and talented *could have* pioneered in the effort to "add Burushaski directly into this [Dene-Caucasic] macrofamily" (as Nikolaev puts it), but so far only Bengtson and Václav Blažek have done so, and so forth.

As pioneers, we have stuck our necks out, and critics have taken their chops. Of course, we have all made mistakes. How could it be otherwise? (If we wanted to be free of criticism, we would have chosen a tiny, obvious language family to work on!)

Nikolaev makes a statement that seems to belittle his own work as well as Chirikba's and mine: "Bengtson's results and Chirikba's results (like our comparison done with Mudrak and to a large extent my 'Dene-Caucasic') are all only interesting claims." Earlier in the same letter he says that "the Sino-Caucasic relationship is already for a long time not just an hypothesis" ... yet ... "the 'Dene-Caucasic' [i.e., Na-Dene + Caucasian, etc.] relationship is not proved, but an hypothesis. Whether I will prove this hypothesis, or someone else, it will be done in time, but for the moment there is a lack of necessary materials ..."

But what is meant by "proof"? Why is there a difference between Nikolaev's assessment of Starostin's hypothesis ("proved") and his own ("not proved" or "only interesting claims"), when they used the same *methods*? Can it be that Nikolaev doubts the adequacy of his Na-Dene reconstructions, and the Na-Dene data they are based on?

It is true that some have been severely critical of Nikolaev 's "Dene-Caucasic" (hereafter "DC") article, but few of the criticisms actually damage the evidence for his case. For example, Michael E. Krauss objects that Nikolaev cited words from the *index* of Krauss' *Eyak Dictionary*, rather than from the entries themselves (see Kaye, 1992). All right, Professor Krauss, suppose we add the full Eyak definitions to "DC". What then, do they strengthen or weaken Nikolaev's case? But Krauss and the others of his school do not do this. Instead, they simply lament that Nikolaev (or Chirikba, or Bengtson, or Greenberg, etc.) has made some mistakes with the language data, and leave it at that. There is no engagement with the real issue the paleo-linguist is concerned with: is x-hypothesis (whether Dene-Caucasic, Amerind, etc.) a viable *hypothesis of classification*? They side-step this issue entirely.

What nonsense! Is this not the eternal plight of the paleo-linguist: those who propose bold new hypotheses are never the same as those who possess a perfect command of *all* the relevant data the critics demand, while those who possess the perfect data never propose bold new hypotheses! In actual fact, you are stuck with ordinary human beings like Nikolaev, Chirikba, and me, who strike boldly and inevitably make some mistakes in the process. The critics forget that mistakes do not negate the good comparisons that remain, and that adding the correct data as often as not strengthens the original comparison.

Let me introduce some examples. Nikolaev gives four etymologies (2.10, 4.3, 5.10, 5.16) that verify the correspondence of Na-Dene *ł* (voiceless lateral fricative) with Caucasic *ł* (the same, noted *λ* by Starostin). One (4.3) compares Eyak *łid* 'deadwood, firewood' with East Caucasic **łwɨndV* 'wood, firewood'. On the face of it, the comparison seems acceptable, phonetically and semantically. Let us see what additional light we can shed on this picture, eventually even bringing in Basque:

I do not own Krauss' *Eyak Dictionary*, nor is it locally available to me. I was able to get it through interlibrary loan for a brief period, during which I checked some of the Eyak words Nikolaev used in comparisons. As regards the Eyak word *ɬid* 'deadwood, firewood', the only relevant additional information I found was that the word had a phonetic variant, *ɬəd*. Blažek and I (LDC p. 23, §72, and addenda p. 162, §72) expanded on the comparison by adding the Eyak *ɬəd* variant, as well as Chipewyan *-ɬìr*, *-ɬə̀r*, *-ɬìr*, *-ɬìy* 'to dry (leaves, bark, grass, etc., in the sun or by fire)' (Chipewyan *-r* from **-d*).

The corresponding East Caucasic words are quite straightforward, e.g. Andi *ɬudi*, Chamalal *ɬunni*, Tsez *redu*, Bezhta *hũdo* (all 'wood, firewood'); Lak (dial.) *x:undu-ti* 'a stick for working leather', etc. (NCED 764). Only in Lak is there a specialized development of the old general meaning.

From Basque, Blažek and I (LDC §72) have introduced **i-ɬinti* 'firebrand, ember, charcoal' (the Basque gloss was unfortunately missing in the published article!), which is attested as *ilhĩ́nti* (the most archaic, a Zuberoan form), *ilinti*, *illindi* (with palatal *ll*), *illeti*, among others in the various Basque dialects. Blažek and I think the old Dene-Caucasic voiceless *ɬ* is preserved in some French Basque dialects as *lh*, but only in medial (intervocalic) position.

In his critique of this comparison, Trask (1995:41) mentions additional meanings: 'dead stalk of wheat; rye grass', etc., noting that the word "has an extraordinary range of meanings ... centering around the notions 'dead', 'black', 'burnt', or 'burning'. It is hard to know what to make of this." Now look back at the Chipewyan words, 'dry (leaves, bark, grass)' (verb). I suggest that all these words revolve around the Paleolithic Dene-Caucasic *fire-making complex*, now attested as various elements of that complex: 'tinder' (dry grass and leaves), 'firewood' (dry wood), leading to 'firebrand, ember, charcoal' as later stages. So Trask's comment, far from "destroying" the etymology, actually strengthens the Dene-Caucasic comparison!

Trask tried further to destroy the comparison by claiming that "the presence of the ... *-nt-* cluster, shows indisputably that this word cannot be ancient." No reason for this is given by Trask, but presumably it relates to the observation, correct as far as it goes, that *-lt-*,*-nt-* usually become *-ld-*, *-nd-* in western Basque dialects (p. 8). I do not think we know enough about Basque phonology yet to discount a comparison this easily. Other aspects of the etymology point to these words being very old indeed. (The form *ilinti*, found in two Bizkaian communities, can be explained either as dialect borrowing, or more likely as a phonetic relic.)

A second example is Nikolaev's (5.10) comparison of Proto-Eyak-Athabaskan **ɬű̄* ~ **ɬű̄w* 'hail, hoarfrost, ice' with East Caucasic **jĭwɬV* / **ɬĭwV* 'snow'. On the Na-Dene side, we have Eyak *ɬaʔ* 'glacier', Hupa *ǩ'i-loo*, *-loy* 'hail', Kato *lō* 'frost', Chipewyan *-lù* 'hard, frozen solid', Navajo *ń-ló* 'hail', Jicarilla *ʔį́-loh* 'ice'. The Caucasic word, with metathetic variants, is found only in Nakh (Chechen *lō*, Batsbi *law*) and Lezgian (Lezgi *žiw*, Tabasaran *jif*) languages (NCED 684).

In our version of the etymology, Blažek and I (LDC 27, §101) add Basque **e-ɬur* 'snow', attested as *élhür* (Zuberoa), *elur* (common Basque), and Bizkaian *erur* and *edur* 'snow'. (The latter two look to me like assimilation and then dissimilation. Trask's [1995:30] supposition that *erur* was the original form cannot be proved.) I now think the Basque word should be analyzed as **e-ɬu-r*, the first element being a fossilized class prefix, the second the root *-ɬu-*, cognate with Na-Dene **ɬu·*, and the third an old plural marker.

Blažek and I have added two comparisons, where however Basque has *initial l-*, where *ɬ* merges with ordinary *l*:

(LDC §19) Basque *lerde* ~ *lirdi* 'drivel, saliva'; East Caucasic **ɬ̱wirdɨ* 'manure, pus': Agul *furd* 'manure', Avar *xʷerd* 'pus', etc. (NCED 763); Sino-Tibetan: Tibetan *lud* 'manure, dung', *lud-pa* 'phlegm, mucus', Kachin *sə-làt* 'sweat', Old Chinese **s-lhi/t/-s* 'snot'; Na-Dene: Athabaskan **ɬu·t'*: Kutchin *ɬid* 'scar', Mattole *-lood-eʔ* 'scar', Hupa *ɬoh*, *-lood-eʔ* 'scab', Chipewyan *ɬùr* 'scab', Navajo *ɬóód*, *-lóód* 'sore'. (The unifying semantic theme is 'bodily secretion/excretion'.)

(LDC §41): Basque **lor̄-* in *lorr-atz* (Bizkaia, Gipuzkoa) 'track, trail, trace, scent'; Caucasic **ɬĕḻħV* 'foot, track': Chechen *lar* 'track, footprint', Tsez *rori* 'foot', etc. (NCED 759); Sino-Tibetan **lăH* 'foot': Chepang *la*, Dimasa *ya*, Miju *p-la* 'foot', Tibetan *b-r-la* 'thigh', etc.; Na-Dene: Eyak *-ɬą·ʔ* 'thigh, upper leg'.

So in sum we have here four etymologies supporting Dene-Caucasic **ł* which have cognates in Basque, Caucasic, and Na-Dene:

a) Basque **i-łinti* 'firebrand', etc.: Cauc **łwindV*: ND **łid*
b) Basque **e-łu-r* 'snow': Cauc **łĭwV*: ND **łu·*
c) Basque *lirdi* 'saliva': Cauc **ł̱wirdɨ*: ND **łu·t'*
d) Basque *lor̄-* 'track': Cauc **łĕḻħV*: ND *łą·ʔ*

The correspondences are entirely regular: voiceless *ł* is reconstructed for Proto-Caucasic and Proto-Na-Dene in all four words, while in Basque the distinction between **ł* and **l* (as well as between other laterals) is maintained only medially, and then only in eastern dialects, notably Zuberoan.

The reasons for repeating the above comparisons are two: (a) to affirm that Nikolaev's correlation between Na-Dene and Caucasic lexics and phonologies is, on the whole, valid, and (b) that the Dene-Caucasic comparisons can, indeed *must*, be extended to include Basque (and Burushaski).

Now that I have defended Nikolaev, I will defend myself (and Blažek and Chirikba) for being so brash as to try to add Basque and Burushaski "directly into this macrofamily" of Dene-Caucasic. We think the greater taxonomic mistake is *not* to add them! What if we did Indo-European using some of the subfamilies, but ignoring Armenian and Greek? But this is the kind of thing Nikolaev (and other Muscovites?) seem to think is correct regarding Dene-Caucasic. Nikolaev seems to think that Basque must be "reconstructed" before it can be compared with Caucasic and the others. Here I am puzzled. Ancient Greek, like Basque, was a language with quite diverse dialects, but no one has ever demanded that "Proto-Greek'' be reconstructed before it could be compared with other Indo-European languages. Usually a dialect (e.g., Ionic, Doric) was selected for use in external comparisons. So with Basque, I think forms from a conservative dialect (e.g., Zuberoan, Bizkaian) should usually be sufficient to use in etymologies.

As to the actual reconstruction of Basque, Nikolaev must not know some of our latest papers, where Blažek and I have made some provisional reconstructions of early Basque. (See, e.g. "Vasco-Caucasic phonological patterns." Bengtson 1995:100-101.) Of course, this kind of reconstruction (based on internal and external evidence) is sometimes incompatible with Trask's or Jacobsen's (based on internal evidence only), so Blažek and I are forced to go it alone.

In defense of Chirikba: he did not just compare Basque with Abkhaz. Many comparisons go beyond Abkhaz to Proto-West Caucasic and even East Caucasic. Of course, some of these we would now call incorrect in view of the *North Caucasian Etymological Dictionary* (NCED, 1994) by Nikola(y)ev & Starostin, but many are still basically correct, in my opinion. Perhaps Chirikba's greatest contribution was recognizing that (North) Caucasic, not Kartvelian, was the closest relative of Basque.

On Nikolaev's estimation that "the true evidence to add Burushaski directly into this macrofamily is lacking," I call attention to my recent paper, "A Comparison of Burushic and (North) Caucasian" (Bengtson 1997), which includes 49 lexical comparisons, a note about the suppletive second person singular pronouns in both families, and the phonetic correspondence of Caucasic lateral affricates (*tl*, *dl*) to Burushic clusters (*lt*, *ld*). This is certainly not the whole story, but I think a "step," one of the steps I refer to below.

Let's not have unrealistic expectations of ourselves and each other as pioneers. Should Nikolaev feel he has to be the Bopp, Rask, Grimm, Schleicher (et al.) of Dene-Caucasic, all rolled into one? Of course not. He has made a beginning (a pretty good beginning, I think), of work that should occupy many other scholars for many years. At what point will that "absolute proof" of DC take place?

My point is that we (Starostin, Nikolaev, Chirikba, Blažek, Ruhlen, and I) are all pioneers in the demonstration of the Dene-Caucasic hypothesis. Let us neither overestimate nor underestimate our importance. Each step, however elementary or even flawed it may seem in hindsight, is necessary. Many others, I hope, will continue our work over many years in search of the "absolute proof" Nikolaev envisions.

Abbreviations

LDC = Lexica Dene-Caucasica: Bengtson & Blažek 1995
NCED = North Caucasian Etymological Dictionary: Nikolaev & Starostin 1994

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

The Genesis of Multilateral Comparison

Joseph H. Greenberg
Stanford University

The present period is one of ferment in historical linguistics, as is evident from the very existence of a journal such as *Mother Tongue*. On the one hand, we have a body of researchers who are presenting evidence for broader groupings and deeper connections than those found in conventional textbooks and, on the other, those who vehemently oppose both their methodology and results.

Discussions regarding methods of classification are now so prominent that the conventional textbooks of historical linguistics and the various surveys of the world's languages are forced to take notice both of methodological issues and concrete new results. These results are perceptibly moving towards a general consensus — what Renfrew has called "the emerging synthesis." In fact, the consensus, as it is now generally realized, is greater than that of the conventional surveys found in historical linguistics textbooks with their numerous doubtful, and undecided cases, when isolated and supposedly unclassifiable languages are considered.

In particular, the appearance of Greenberg (1987) has already generated a sizeable controversial literature. Much of this has centered on a critique of the notion of mass comparison first stated in Greenberg (1954) and subsequently renamed "multilateral comparison". The reason for this merely terminological charge was that the term "mass comparison" conjures up a vision of crudity, whereas the method itself is both a subtle and powerful one. Indeed, when confronted with an area containing a vast number of apparently disparate languages such as Africa, New Guinea, or the Americas, it is difficult to see how one could start otherwise.

One strange aspect of the present controversy is that, although it was first stated more than forty years ago and its results in Africa accepted, there was virtually no discussion of it. The more comprehensive statement in *Essays in Linguistics* (1957) indeed formed the basis of discussion for the methodology of classification in the large compilation on American Indian languages edited by Campbell and Mithun (1967). It is apparent however, that, though cited as authoritative, the methodology was not well understood, as shown by its posing the question in terms of relationship rather than classification. It has become clear that it was not the method itself, but the unexpected and unwelcome results obtained in the Americas by its application that provided the impetus for the present controversy.

To begin with, it should be made clear that this controversy has nothing to do intellectually with the opposition between "splitters" and "lumpers," or as Matisoff stated it, microlinguists versus macrolinguists. In biology, the most inveterate splitter believes in the reality of vast assemblages of species with such enormous phenetic differences as bats and whales (both mammals) or for that matter at a deeper level horses and trees (both eukaryotes)! What all historical linguists should view as fundamental is a taxonomy of valid genetic groups of different levels. Traditionally, indeed, historical linguistics has been based on the studies of stocks, families, etc., each implying a single linguistic ancestor, a "proto-language." The size of a family is irrelevant, it is the number and types of resemblances among the languages that counts.

My earliest attempt at African classification (1948-49) resulted in sixteen families, and I only later successively reduced that to twelve and to four in order to account for facts of differential resemblance among groups that would otherwise be unaccounted for.

In my earliest work on the Americas, which began in South America, I thought of seven independent families. It was as certain facts began to obtrude themselves (and the now famous *n* first person, *m* second person was but one of them), that I began to formulate the Amerind hypothesis. This was also the case with Indo-Pacific.

I should add that all this has nothing to do with scientific caution either. Caution has to do with a relation between evidence and conclusion. It is just as wrong to refuse to draw a conclusion that would account for certain facts as to draw incorrect conclusions from insufficient evidence, or as is more frequently the case, evidence pointing to a different classification.

The present essay arises from a suggestion by Hal Fleming about how I came to formulate the notion of multilateral (*né* mass) comparison. It turns out that introspection is a treacherous guide to events of fifty or so years ago and that research into my own writings and those of others reveals a more complex story than I had first imagined from the distorted residue of half century old events. In doing this, I was assisted by George Grace, whose graduate advisor I had been in the Columbia Anthropology Department. He started in 1949 and received his Ph.D. in 1958. During this period, he compiled comparative notebooks for me on Central Saharan, Central Sudanic, and

Eastern Sudanic, all of which I still have, with numerous later additions in my own hand. Unfortunately, I neither kept a diary nor dated my notebooks. George's recollections were invaluable in reconstructing the events of this period.

As noted above, when I responded to Hal Fleming's suggestion, I thought the story would be a straightforward one. In fact, two major problems arose in reconstructing this narrative. To begin with, I enunciated not one but three major principles in regard to classification, and of these, multilateral comparison was the latest to be stated, namely, in the introduction to the eighth and last article on African linguistic classification in the *Southwestern Journal of Anthropology*_(1954). The first principle was stated in the first paragraph of the first article (1948). This was the resolute exclusion of typological resemblances as relevant to classification. The reason was that, in the literature on African classification up to that point, the use of such criteria, except at the lowest levels, was all-pervasive.

I was well aware from my knowledge of historical linguistics that such typological resemblances were far less stable than the cumulative weight of numerous resemblances in sound-meaning, both morphological and lexical. For example, great emphasis was placed on the presence or absence of sex gender in the literature, yet it would wreak havoc if applied in better known areas. Thus French would be classified with Hebrew because it had two genders based on sex, but would be separated from the Indo-European Armenian which had no gender. We may say that these and other similar erroneous notions all involved what might be called the "thesis of discontinuity." That is, while nobody applied such criteria, which might on occasion even separate two dialects of the same language, when the resemblances were overwhelming and obvious, as soon as scholars went beyond these obvious resemblances, they embraced apparently "deeper" structural criteria of known historical fragility. Such criteria are still used in some parts of the world.

The second principle was first stated in 1949 in the fifth article of the series, namely, that languages should be classified on the basis of linguistic evidence alone. When presented in this form, it is so obvious that it would hardly seem to need stating, but it was constantly violated in African classification and continues to be violated in other areas today.

The immediate occasion for stating it was the question of Nilotic, generally separated from "Nilo-Hamitic" within the East Sudanic grouping I postulated. Objectively, one had three groups: (1) Nilotic (Shilluk, Nuer, etc.), (2) part of "Nilo-Hamitic" (Masai, Turkana, etc.), and (3) a southern group (Nandi-Suk) also called "Nilo-Hamitic." A classification excluding extraneous non-linguistic evidence would arrive at the conclusion that the three formed a single subgroup with Eastern Sudanic and that groups (1) and (2) were more closely related to each other than either was to (3). However, (2) and (3) at that time were generally grouped together as a "Nilo-Hamitic" hybrid, or by some as Hamitic, based on linguistic gender, a culture centered on cattle, and supposed Hamitic physical features. Other examples of the violation of this principle were the separate status of Semitic within Hamito-Semitic because of its cultural importance and the special treatment accorded Bantu because of its enormous area and population.

The third principle, mass comparison, came later in the eighth article of the series, following the previous ones by four years (1954:406). Unlike the first two principles, this did not arise from noting what seemed to me obvious errors on the part of previous classifiers, but because of certain questions raised by critics. Those named there are Mary Haas in conversation and McQuown in his review in the *American Anthropologist*, of the volume *Anthropology Today* (1953) edited by Kroeber, which contained my article "Historical Linguistics and Unwritten Languages." They raised two major points, the feasibility of drawing genetic conclusions of broad scope and the supposed arbitrariness of the results, in that just as good evidence might be assembled for other classifications cutting across the families I had proposed.

In thinking about these objections, I realized that there was a principle that had been unstated, because it seemed too obvious to me, namely, that one assembled as broad a data base as possible and included all of it in one's comparisons. From such a data base, coherent groupings appeared almost immediately and were confirmed again and again. This was evidently not the way that these other scholars proceeded. Comparisons were piecemeal, resulting in the ultimate absurdity of binary, contextless comparisons. Moreover, there was a strong inhibition against using earlier sources which were imperfectly recorded. Hence, I stated "The importance of mass comparison as opposed to isolated comparisons between pairs of languages has become clear to me as a result of certain questions of a general nature raised by my critics." It is clear that the critics have never actually tried the methods I advocated. This applies to Campbell's "fish in a barrel" analogy repeated again and again in the present day literature. Supposedly, as the number of languages increases, the number of possible pairs of languages increase and with it the number of resemblances so that you can prove anything.

I believe a sensible analysis will reveal the fallaciousness of such an assertion. It advances the paradoxical notion that the broader the base of relevant data, the more uncertain the conclusions. Of course, the number of two-way resemblances increases, but what one should be looking for as new languages are added is the appearance of widespread and manifold, not two-way resemblances characteristic of already distinguished groups. So Hrozný, credited with the discoveries that convinced Indo-Europeanists of the affiliation of Hittite, did not compare Hittite with say, Slavic, to show they are related, but pointed to concrete morphological lexical resemblances each found in a number of Indo-European branches.

The article in which the notion of multilateral comparison was first enunciated (1954) reduced the number of African language families from sixteen to twelve, but not to the four of now generally accepted classification. The fourfold classification first appears in the article "African Languages" in *Collier's Encyclopedia* (1959) and in the popular article "African Languages and Tribes" (1960) in the *Rotarian*. All this was finally presented in detail in *Languages in Africa*_(1963). There was thus a hiatus of about four years (1954-59) in which, although the principle of multilateral comparison had been stated, it was not applied in Africa.

The reason was that, although the classification into twelve independent African groups was not satisfactory, the question was put on hold because of other interests. One was my work on typology and universals. The other was that I had already begun work on the classification of two other areas. As early as 1951, work began on South America and was soon extended to North America. One by-product was the discovery that Jicaque of Honduras, hitherto considered a very problematic case, was Hokan. This resulted in a joint paper with Swadesh (1953). The proposed classification of Jicaque as Hokan is now generally accepted. In 1956, in a paper read at the Congress of Anthropological and Ethnological Science at Philadelphia, I outlined a three-fold classification of the languages of the Americas without subgrouping details north of Central America. At the December 1951 meeting of the Linguistic Society of America, I gave a paper called "The Classification of Australian Languages," the results of which were never published in detail. I outlined them in the paper "Historical Linguistics and Unwritten Languages" in *Anthropology Today* edited by Kroeber. Although only briefly outlined there, it is clear that these results are essentially consonant with those held today. For "General Australian," substitute Pama-Nyungan, and the basic outlines are virtually identical.

I was largely stimulated by these extra-African endeavors and by Kroeber's presence at Columbia. The question of Africa, had, as we have seen, been tabled for a while, but my experience with American and Australian data in this interim period provided me with additional experience and perspective.

I remember, as one of those transforming moments that one experiences only a few times in one's life, that one day, probably in early 1959, as I put my foot on the pavement to cross Amsterdam Avenue on my way to Columbia, an idea flashed before me. Why shouldn't I just look at all of my then twelve families in Africa together? Nothing changes methodologically just because the groupings are more distant from each other. I had, already, to a great extent with George Grace's previous help, all the necessary material in my notebooks. I rapidly arrived at the conclusion that of the smaller groups distinct from Niger-Congo, Afroasiatic, Khoisan, and Macro-Sudanic, Kordofanian belonged with Niger-Congo and the rest with Macro-Sudanic. One Kordofanian group which I call Tumtum, and which is now called Kadugli-Krongo, seemed somewhat separate from the rest of Kordofanian. In fact, I said so in *Languages of Africa* (1963:149) noting that "the fourth (Tumtum) shows considerable divergence from the remainder." I remember speculating that it might be a separate stock related to both Niger-Congo and Nilo-Saharan. However, some of the resemblances in paired singular plural suffixes, notably *t/n*, *s/ny*, and *b/y*, persuaded me that it probably belonged with Kordofanian.

Finally, I should like to point out, because of its parallel to the present Amerind situation, that, at what was still the major center of African linguistic studies in the world, the Department of Africa at the School of Oriental and African Studies in London, bitter opposition to me prevailed among the linguists, with a few covert exceptions. Just as I now draw support from geneticists and archaeologists, it was non-linguists, in this case, historians and archaeologists interested in Bantu origins, who finally saw the superiority of my explanations and who ultimately turned the tide. In several interesting articles, Colin Flight, a British historian at the School, recounted the struggles within the institution. The attitude of British linguists at that time was so similar to that of most American linguists now, that it is worth concluding by quoting Flight at some length (1981:91):

> It was British Africanists who expressed the most strenuous opposition to Greenberg's classification. In short the reaction from British Africanists was almost uniformly unconstructive. Their attitude at first was amused indifference. Privately, they felt, publicly they were committed,

> to the proposition that classification was a task of extreme difficulty. Anyone who thought it was easy did not even deserve to be taken seriously. Even by making it *look* easy, however, Greenberg was threatening their authority. It was clear that he aroused deep feelings of resentment which were betrayed by a remarkable propensity for misunderstanding and misrepresenting what Greenberg had to say.

And again (1988:266) "Rather than trying to understand what Greenberg had to say, British linguists seem to have done their best to misunderstand... In a word, they acted like lawyers."

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue II (December 1996)

References for *Mother Tongue II*

Nihali Section

Andronov, M. S. 1979. *Sravnitel'naja grammatika dravidijskix jazykov*. Moscow: Nauka.

Batchelor, John. 1938. *An Ainu-English-Japanese Dictionary*. Tokyo: Iwanami.

Bengtson, John D. 1992. "A Case for the Austric Affiliation of Ainu" (one-page abstract), in Shevoroshkin (ed.) 1992, p. 364.

Bengtson, John D. 1994. "Macro-Australic". *Mother Tongue* (newsletter) 23:73-75.

Bengtson, John D. and Václav Blažek. 1996. "Lexical Parallels between Ainu and Austric". Ms.

Berger, Hermann. 1959. "Deutung einiger alter Stammesnamen der Bhil aus der vorarischen Mythologie des Epos und der Purana". *Wiener Zeitschrift zur Kunde der Südost-Asien* 3:34-82.

Bhattacharya, Sudhibhushan. 1957. "Field Notes on Nahali". *Indian Linguistics* 17:245-258.

Bhattacharya, Sudhibhushan. 1966. "Some Munda Etymologies", in Zide (ed.) 1966, pp. 28-60.

Bhattacharya, Sudhibhushan. 1977. "The Munda Languages in Ancient India". *Journal of Indian History* 55:25-33.

Blažek, Václav. 1992. "Kartvelian Material in Nostratic Lexicon: New Etymologies", in Shevoroshkin (ed.) 1992, pp. 112-128.

Burrow, Thomas. 1958. "Sanskrit and the Pre-Aryan Tribes and Languages". *Bulletin of the Ramakrishna Institute of Culture*. (February 1958.)

Burrow, Thomas. 1963. Review of Kuiper 1962. *Bulletin of the School of Oriental and Asiatic Studies* 26:436-437.

Burrow, Thomas and Murray B. Emeneau. 1984. *A Dravidian Etymological Dictionary*. 2nd edition. Oxford: Clarendon Press.

Chiri, Mashihi. 1966. *Ainugo nyuumon [An Introduction to the Ainu Language]*. Sapporo: Hokkaido shuppan kiga sentaa.

De Candolle, B. 1961. "Contribution à l'étude des Nahals". *Anthropos* 56:750-788.

Dolgopolsky, A. B. 1973. *Sravnitel'no-istoričeskaja fonetika kušitskix jazykov*. Moscow: SFKJ.

Dolgopolsky, A. B. 1984. "On Personal Pronouns in the Nostratic Languages", in *Linguistica et Philologica: Gedenkschrift für Björn Collinder*, ed. By. O. Gschwantler, K. Rédei, and H. Reichert, pp. 65-112. Vienna: Braumüller.

Efimov, A. J. 1990. *Istoričeskaja fonologija južnobahnaričeskix jazykov*. Moscow: Nauka.

Forsyth, J. 1870. *Report of the Revenue Settlement of British Nimar District of the Central Provinces*.

Forsyth, J. 1872. *The Highlands of Central India*. London.

Fuchs, S. 1963. "The Aboriginal Races in North-Eastern Central India", in *Festschrift Paul Schebesta (Studi Instituti Anthropos* 18), pp. 349-355. Wien-Modling: St. Gabriel Verlag.

Fuchs, S. 1988. *The Korkus of the Vindhya Hills*. New Delhi.

Gjerdman, Olof. 1926. "Word parallels between Ainu and other languages." *Le Monde Oriental*_20:29-84.

Grierson, George (ed.). 1906. *Linguistic Survey of India*, vol. 4. (The article on Nihali, prepared by Sten Konow, is on pp. 185-189.)

Guiraud, P. 1956. *L'Argot*. (*Que Sais-je*, no. 700.) Paris: Presses Universitaires de France.

Hayes, LaVaughn H. 1993. "On the Track of Austric: Part I. Introduction." *Mon-Khmer Studies* 21:143-177.

Hermanns, M. 1966. *Die religiöse-magische Weltanschauung der Primitivstamme Indiens*. (Band 2: Die Bhilala, Korku, Gond, Baiga). Wiesbaden: Franz Steiner Verlag.

Illič-Svityč, V. M. 1967. "Materialy k sravnitel'nomu slovarju nostratičeskix jazykov," in *Etimologija* 1965, ed.. by O. Trubačev, et al., pp. 321-373. Moscow..

Illič-Svityč, V. M. 1971. *Opyt sravnenija nostratičeskix jazykov. Sravnitel'nyj slovar'* (*b-ḳ*). Moscow: Nauka.

Illič-Svityč, V. M. 1976. *Opyt sravnenija nostratičeskix jazykov. Sravnitel'nyj slovar' (l-ʒ́). Ukazateli.*_Moscow: Nauka.

Jungraithmayr, H. and D. Ibriszimow. 1994. *Chadic Lexical Roots*. 2 vols. Berlin.

Klein, E. 1987. *A Comprehensive Etymological Dictionary of the Hebrew Language*. New York, London: Macmillan.

Klimov, G. A. 1964. *Etimologičeskij slovar' kartvel'skix jazykov*. Moscow: Nauka.
Konow, Sten. 1906.: see Grierson (ed.) 1906.
Konow, Sten. 1908. "Notes on the Munda family of speech in India." *Anthropos* 3:68-82.
Koppers, W. 1947. *Geheimnisse des Dschungels*. Lucerne: Verlag Josef Stocker.
Koppers, W. 1948. *Die Bhil in Zentralindien*. Horn-Wien: Verlag Ferdinand Berger.
Kuiper, F. B. J. 1962. "Nahali: A Comparative Study," in *Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afd. Letterkunde*, (Nieuwe Reeks) 25/5, pp. 239-352. Amsterdam.
Kuiper, F. B. J. 1966. " The sources of the Nahali vocabulary," in Zide (ed.) 1966, pp. 57-81.
Kuiper, F. B. J. 1967. "The Genesis of a Linguistic Area." *Indo-Iranian Journal* 10 (2/3).
Kuiper, F. B. J. 1972. Review of *Current Trends in Linguistics*. *Indo-Iranian Journal* 14 (3/4). (See pp. 292-293.)
Lynch, Owen. 1963 (?). "Outline of Nihali Grammar." Unpublished typescript.
Mehrotra, R. R. 1977. *Sociology of Secret Languages*. Simla: Indian Institute of Advanced Study.
Menges, Karl-Heinrich. 1960. *Morphologische Probleme I. Zum Genetiv und Akkusativ*. Wiesbaden: Harrassowitz.
Mufwene, S. 1995. "Jargons, pidgins, creoles, koines: what are they?," in *Pidgins and creoles: structures and status*, ed. by D. Spears and D. Winford. Amsterdam: John Benjamins.
Mundlay, Asha. 1964 (revised 1972). "Nihali (Kalṭu-manḍi): Its place among the Munda languages of India." Unpublished typescript.
Mundlay, Asha. 1979 (1997). "Nihali Lexicon." (In this volume.)
Mundlay, Asha. 1995. *The Nihals* (in Marathi).
Murayama, Shichiroo. 1966. "Mongolisch und Japanisch. Ein Versuch zum lexikalischen Vergleich," in *Collectanea Mongolica, Festschrift Rintchen* (Asiatische Forschungen 17), pp. 153-156. Wiesbaden: Harrassowitz.
Murayama, Shichiroo. 1992a. "Ein Versuch der Etymologisierung einiger Ainu-Wörter," in *Nenrin — Jahresringe* (Festgabe für Hans A. Dettmer), ed. by K. Müller and W. Naumann, pp. 148-157. Wiesbaden: Harrassowitz.
Murayama, Shichiroo. 1992b. *Ainugo no kigen* [Origins of the Ainu Language]. Tokyo: San'ichi shoboo.
Murayama, Shichiroo. 1993. *Ainugo no kenkyuu* [A Study of the Ainu Language]. Tokyo: San'ichi shoboo.
Murayama, Shichiroo. 1995. *Nihongo no hikaku kenkyuu* [A Comparative Study of the Japanese Language]. Tokyo: San 'ichi shoboo.
Newman, Paul. 1977. "Chadic classification and reconstructions." *Afroasiatic Linguistics* 5/1:1-42.
Orël, V.E. and O.V. Stolbova. 1995. *Hamito-Semitic Etymological Dictionary*. Leiden, New York, Köln.
Peiros, Ilia. 1992. "The Austric Macrofamily: Some Considerations," in Shevoroshkin (ed.) 1992, pp. 354-363.
Piłsudski, Bronislaw. 1912. *Materials for the Study of the Ainu Language and Folklore*. Krakow.
Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Harrassowitz.
Pinnow, Heinz-Jürgen. 1963. "The position of the Munda languages within the Austroasiatic Language Family," in *Linguistic Comparison in South East Asia and the Pacific*, ed. by H. L. Shorto, pp. 140-152. London: School of African and Oriental Studies.
Pinnow, Heinz-Jürgen. 1965. "Personal pronouns in the Austroasiatic languages: a historical study," in *Indo-Pacific Linguistic Studies*, Part I, pp. 3-42. Amsterdam.
Pinnow, Heinz-Jürgen. 1966a. Review of Kuiper 1962. *Orientalische Literaturzeitung* 61:492-497.
Pinnow, Heinz-Jürgen. 1966b. "A comparative study of the verb in Munda languages," in Zide (ed.) 1966, pp. 96-193.
Rédei, Károly. 1988. *Uralisches etymologisches Wörterbuch*. Budapest: Kiado.
Russell, R. J. and Hira Lal. 1916. *Tribes and Castes of the Central Provinces*, vol. 4 (Nihals on pp. 259-261).
Shafer, Robert. 1940. "Nahali. A Linguistic Study in Paleo-ethnography". *Harvard Journal of Asiatic Studies* 5:346-371.
Shafer, Robert. 1954. *Ethnography of Ancient India*. Wiesbaden: Harrassowitz.
Shevoroshkin, Vitaly (ed.). 1992. *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum: Brockmeyer.
Sidwell, Paul. 1996. "The External Relations of Ainu: Problems and Prospects." Ms.
Singh, K. S. and S. Manoharan. 1993. *People of India, National Series* (vol 9, Languages and Scripts). Oxford, Delhi: Anthropological Survey of India.
Skeat, W. W. and C. O. Blagden. 1906. *Pagan Races of the Malay Peninsula*. London: Macmillan.
Sternberg, Lev [Leo]. 1929. "The Ainu Problem." *Anthropos* 24:755-799.

Sternberg, Lev [Leo]. 1933. *Giljaki, Oroči, Gol'dy, Negidal'cy, Ainy* [Gilyaks, Orochs, Golds, Neghidals, Ainus]. Khabarovsk: Dal'giz.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Thompson, Charles S. 1895. *Rudiments of the Bhili Language*. Ahmedabad.

Vovin, Alexander. 1992. "The Origins of the Ainu Language," in *Pan-Asiatic Linguistics* (Proceedings of the Third International Symposium on Language and Linguistics), vol. 2, pp. 672-685.

Vovin, Alexander. 1993. *A Reconstruction of Proto-Ainu*. Leiden: E. J. Brill.

Zide, Norman H. (ed.). 1966. *Studies in Comparative Austroasiatic Linguistics*. The Hague: Mouton.

Zide, Norman H. 1968. "Some comparative notes on Gtaʔ pronouns," in *Pratidanam*, pp. 348-358. The Hague: Mouton.

Zide, Norman H. 1969. "Munda," in *Current Trends in Linguistics*, vol. 5. The Hague: Mouton.

Zide, Norman H. 1972. "Some Munda Etymological Notes on Names in the Ramayana." *Journal of the Ganganatha Jha Kendriya Sanskrit Vidyapeeth* 28:755.

Zide, Norman H. 1978. *Studies in the Munda Numerals*. Mysore: Central Institute of Indian Languages.

Other Papers

Bengtson, John D. 1991. "Macro-Caucasian Phonology (Part I)". In Shevoroshkin 1991, pp. 142-149.

Bengtson, John D. 1992. "Macro-Caucasian Phonology (Revised Version)". In Shevoroshkin 1992, pp. 342-351.

Bengtson, John D. 1994. "On the Genetic Classification of Basque". *Mother Tongue* (Newsletter) 22:31-36.

Bengtston, John D. 1995. "Basque: An Orphan Forever? A Response to Trask". *Mother Tongue* (Journal) I:84-103.

Bengtson, John D. 1997. "Ein Vergleich von Buruschaski und Nordkaukasisch." (To appear in *Georgica*.)

Bengtson, John D. and Václav Blažek. 1995. "Lexica Dene-Caucasica." *Central Asiatic Journal* 39/1:11-50, 39/2: 161-164.

Blažek, Václav. 1995. "Towards the Position of Basque: A Reply to Trask's Critique of the Dene-Caucasian Hypothesis". *Mother Tongue* (Journal) I:104-110.

Blažek, Václav and John D. Bengtson. 1995. "Lexica Dene-Caucasica". *Central Asiatic Journal* 39/1:11-50; 39/2:161-164.

Campbell, Lyle, and Marianne Mithun (eds.). 1979. *The Languages of Native America*. Austin.

Corominas, Joan and José A. Pascual. 1980. *Diccionario crítico etimológico castellano e hispánico*. 6 vols. Madrid: Gredos.

Diakonoff, Igor M. and Sergei A. Starostin. 1986. *Hurro-Urartian as an Eastern Caucasian Language*. (= *Münchener Studien zur Sprachwissenschaft*, Beiheft 12.) Munich: Kitzinger.

Flight, Colin. 1981. "Trees and Traps: Strategies for the Classification of African Languages". *History in Africa* 8:43-74.

Greenberg, Joseph H. 1949-1950, 1954. "Studies in African Linguistic Classification". *Southwestern Journal of Anthropology* 5, 6, 10.

Greenberg, Joseph H. 1957. *Essays in Linguistics*. Chicago: University of Chicago Press.

Greenberg, Joseph H. 1960 (1956). "The General Classification of Central and South American Languages". In *Selected Papers of the Fifth International Congress of Anthopological and Ethnological Sciences*, ed. by A. Wallace, pp. 791-794. Philadelphia: University of Philadelphia Press.

Greenberg, Joseph H. 1963. *The Languages of Africa*. Bloomington: University of Indiana.

Kaye, A. S. 1992. "Distant Genetic Relationship and Edward Sapir". *Semiotica* 91.3/4:273-300.

Nikolaev, S. L. 1991. "Sino-Caucasian Languages in America". In Shevoroshkin 1991, pp. 42-66.

Nikolayev, Sergei L. and Sergei A. Starostin. 1994. *A North Caucasian Etymological Dictionary*. Moscow: Asterisk Press.

Ruhlen, Merritt. 1994. *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.

Sarich, Vincent M. 1994. "Occam's Razor and Historical Linguistics". In *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. by Matthew Y. Chen and Ovid J. L. Tzeng, pp. 409-430. Taipei: Pyramid Press.

Shevoroshkin, Vitaly (ed.). 1991. *Dene-Sino-Caucasian Languages*. Bochum: Brockmeyer.

Shevoroshkin, Vitaly (ed.). 1992. *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum: Brockmeyer.

Starostin, Sergei A. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy*. Moscow: Nauka.

Starostin, Sergei A. 1991. "On the Hypothesis of a Genetic Connection between the Sino-Tibetan Languages and the Yeniseian and North Caucasian Languages". In Shevoroshkin 1991, pp. 12-41.

Starostin, Sergei A. 1995. *Sravnitel'nij slovar' enisejskix jazykov* [Comparative Dictionary of the Yeniseian Languages]. In *Ketskij Sbornik, Lingvistika* (Studia Ketica, Linguistics), ed. by S. A. Starostin, pp. 176-313. Moskow: Škola Jazyki Russkoj Kul'tury.

Starostin, Sergei A. and Merritt Ruhlen. 1994. "Proto-Yeniseian Reconstructions, with extra-Yeniseian Comparisons". In Ruhlen 1994, pp. 70-82.

Trask, R. L. 1994-1995. "Basque: The Search for Relatives (Part 1)". *Dhumbadji!* 2/1:3-54.

Trask, R. L. 1995. "Basque and Dene-Caucasian: A Critique from the Basque Side". *Mother Tongue* (Journal) I:3-82.

Trask, R. L. 1996. *The History of Basque*. London: Routledge.

Abbreviations

Arch.	Archi (North Caucasic language)
Bq	Basque (Euskara)
LDC	"Lexica Dene-Caucasica" (= Blažek and Bengtson 1995)
Lezg.	Lezgi (North Caucasic language)
NCED	*North Caucasian Etymological Dictionary* (= Nikolayev and Starostin 1994)
OCh.	Old Chinese
PEC	Proto-East Caucasic
PNC	Proto-North Caucasic
PST	Proto-Sino-Tibetan
PY	Proto-Yeniseian

MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY

AIM AND SCOPE

The Association for the Study of Language in Prehistory (ASLIP) is a nonprofit organization, incorporated under the laws of the Commonwealth of Massachusetts. Its purpose is to encourage and support the study of language in prehistory in all fields and by all means, including research on the early evolution of human language, supporting conferences, setting up a databank, and publishing a newsletter and a journal to report these activities.

MEMBERSHIP

Annual dues for ASLIP membership and subscription to *Mother Tongue* are US $25.00 in all countries except those with currency problems. For membership information, contact:

Harold C. Fleming, *Secretary-Treasurer*
Association for the Study of Language in Prehistory
16 Butman Avenue
Gloucester, MA 01930-1006 U.S.A.

Design and Production:

SIGNUM
DESKTOP PUBLISHING
88-B QUEEN STREET
CHARLESTON, SC 29401-2427
(803) 720-8531 (PHONE & FAX)